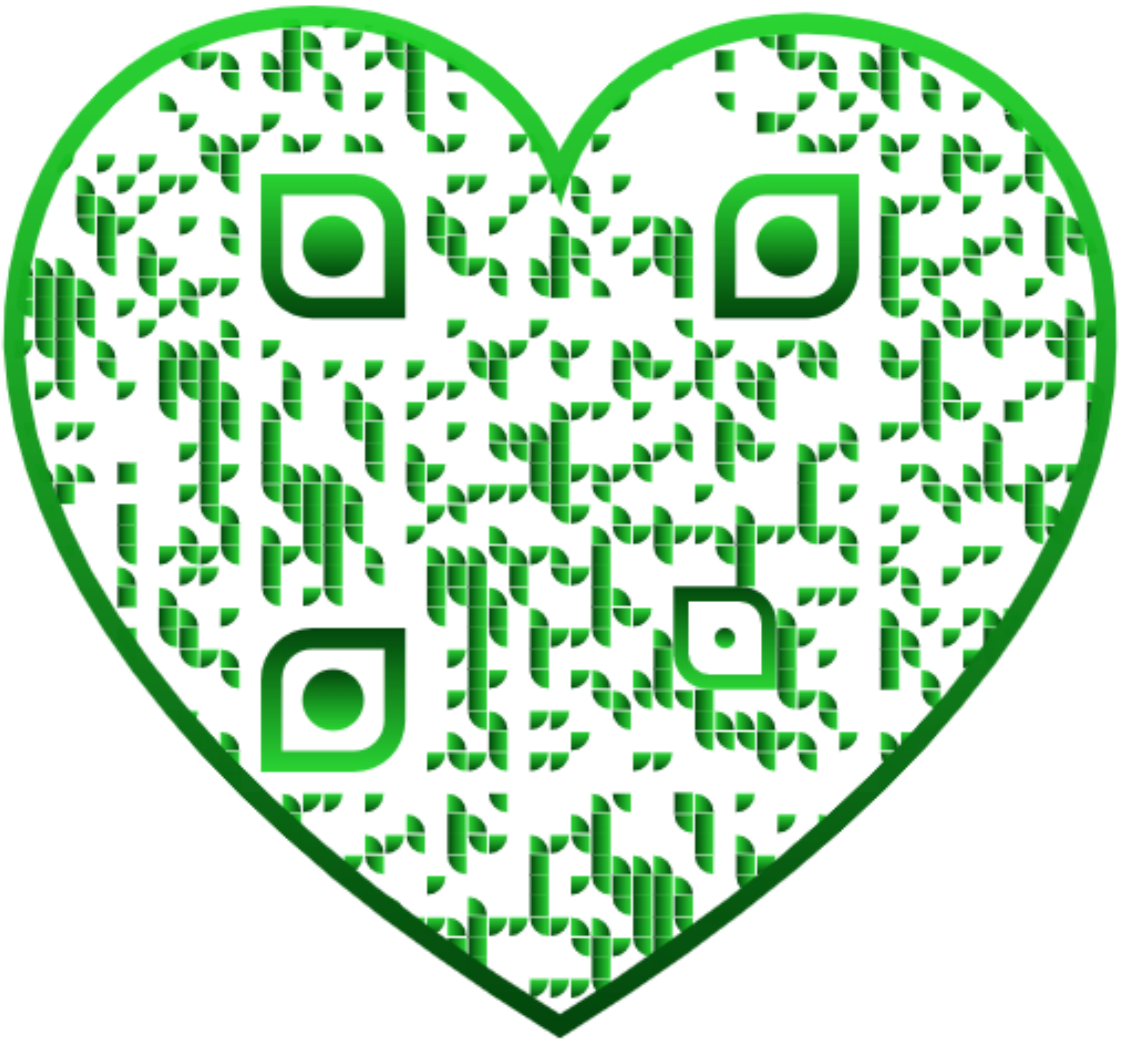
Philippe Caldero
et Jérôme Germoni

Histoires hédonistes de groupes et de géométries

Tome premier

Calvage & Mounet



101. — Jacques Faraut. *Analyse sur les groupes de Lie. Une introduction*
102. — Patrice Tauvel. *Corps commutatifs et théorie de Galois*
103. — Jean Saint Raymond. *Topologie, calcul différentiel et variable complexe*
104. — Clément de Seguins Pazzis. *Invitation aux formes quadratiques*
105. — Bruno Ingrao. *Coniques projectives, affines et métriques*
106. — Wolfgang Bertram. *Calcul différentiel topologique élémentaire*
107. — Henri Lombardi et Claude Quitté. *Algèbre commutative. Méthodes constructives*
108. — Frédéric Testard. *Analyse mathématique. La maîtrise de l'implicite*
109. — Grégory Berhuy. *Modules : théorie, pratique... et un peu d'arithmétique*
110. — Bernard Candelpergher. *Théorie des probabilités. Une introduction élémentaire*
111. — Philippe Caldero et Jérôme Germoni. *Histoires hédonistes de groupes et de géométries. Deux tomes.*

Philippe Caldero et Jérôme Germoni

Histoires hédonistes de groupes et de géométries

Tome premier

Calvage & Mounet

Philippe Caldero est ancien élève de l'ÉNS de Saint-Cloud, agrégé de mathématiques et docteur. Il est maître de conférences à l'université Lyon 1 et ses travaux de recherche concernent la théorie des représentations.

Jérôme Germoni est ancien élève de l'ÉNS, agrégé de mathématiques et docteur. Maître de conférences à Lyon 1. Passionné de questions d'enseignement et de diffusion de la culture mathématique, il a dirigé l'Institut de recherche sur l'enseignement des mathématiques (IREM) de Lyon.

germoni@math.univ-lyon1.fr
caldero@math.univ-lyon1.fr

Mathematics Subject Classification (2000) :

14-XX Algebraic geometry
  14H-XX Curves
    14.20 Algebraic curves, surfaces and special varieties
51-XX Geometry
  51F-XX Metric geometry
  51N-XX Analytic and descriptive geometry
    51N10 Affine analytic geometry
    51N15 Projective analytic geometry
    51N20 Euclidean analytic geometry
    51N25 Analytic geometry with other transformation groups
    51N30 Geometry of classical groups
    51A05 General theory and projective geometries

ISBN 978-2-91-635231-2

♾ Imprimé sur papier permanent

© Calvage & Mounet, Paris, 2013

À tous ceux qui nous ont fait découvrir les plaisirs mathématiques :
à nos parents, nos professeurs...
qui sont parfois un peu les mêmes...

Table des matières

I. Actions et théorèmes du rang

1. Théorème du rang 2
2. Action de $\mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$ sur $\mathscr{M}_{m,n}(\mathbb{K})$ par équivalence . . . 3
3. Stabilisateur, quotient et orbite 5
4. Topologie des orbites ($\mathbb{K}$ corps des réels ou des complexes) . . . 9

A. Annexe. Actions de groupes 11
B. Développement. Algorithme de Gauss 18
C. Exercices 21

II. Groupes topologiques, actions continues

1. Groupes topologiques 26
2. Topologie et algèbre linéaire 28
3. Action continue, théorème d'homéomorphisme 31
4. Applications du théorème d'homéomorphisme 37
5. Produits semi-directs topologiques 43

A. Annexe. Normes sur les espaces de matrices 51
B. Annexe. Connexité 53
C. Annexe. Compacité locale 55
D. Annexe. Topologie sur un espace quotient 58
E. Développement. Actions classiques et leurs invariants 62
F. Exercices 65

III. Réduction des endomorphismes

1. Action $\mathrm{GL}_n(\mathbb{C})$ sur $\mathscr{D}_n(\mathbb{C})$ par conjugaison 84
2. Action $\mathrm{GL}_n(\mathbb{C})$ sur $\mathscr{N}_n(\mathbb{C})$ par conjugaison 87
3. Action $\mathrm{GL}_n(\mathbb{C})$ sur $\mathscr{M}_n(\mathbb{C})$ par conjugaison 98
4. Passage de $\mathbb{C}$ à $\mathbb{R}$ 100
5. Invariants de similitude sur un corps quelconque 101
6. Invariants de similitude et diagrammes de Young 108

A. Annexe. Lemme des noyaux et décomposition de Dunford 110
B. Annexe. Partitions 113
C. Développement. Partitions de 6 118
D. Exercices 120

IV. Matrices échelonnées et grassmanniennes

1. Motivation 128
2. Matrices échelonnées en lignes 130
3. Décomposition en cellules des grassmanniennes 135
A. Annexe. Opérations élémentaires sur les rangées 143
B. Exercices 144

V. Groupes conservant une forme bilinéaire

1. Classification des formes quadratiques 149
2. Le cas hermitien 156
3. Les groupes orthogonaux et leurs actions 157
4. Formes bilinéaires antisymétriques 160
5. Lien avec la réduction 162
6. Classification des coniques 164
A. Annexe. Généralités sur les formes bilinéaires 170
B. Annexe. Théorie classique des formes quadratiques 175
C. Développement. Réciprocité quadratique 182
D. Exercices 187

VI. Décomposition polaire et applications

1. Théorème de décomposition polaire 201
2. L'exponentielle 207
A. Décomposition polaire pour $\mathrm{O}(p, q)$ 210
B. Exercices 216

VII. Le corps des quaternions

1. Le corps des quaternions 224
2. Interlude: l'ours mal peigné 230
3. Applications à $\mathrm{SO}(3)$ 232
4. Applications à $\mathrm{SO}(4)$ 234
5. Aparté sur la simplicité 235
A. Annexe. Engendrement de $\mathrm{O}(n)$ et $\mathrm{SO}(n)$, simplicité de $\mathrm{SO}(3)$. 237
B. Développement. Fibration de Hopf 240
C. Exercices 242

VIII. Combinatoire algébrique

1. Dénombrement sur les corps finis 250
2. Décomposition en cellules et dénombrement 255
3. Applications aux isomorphismes exceptionnels de groupes finis . 256
4. Développement. Un non-isomorphisme exceptionnel 262
A. Exercices 263

IX. Groupes de Lie
1. Groupes de Lie classiques 268
2. Applications aux isomorphismes exceptionnels 272
A. Annexe. Rappels et compléments en calcul différentiel 275
B. Notion d'algèbre de Lie 283
C. Exercices 286

X. Droite projective et applications
1. Droite projective sur un corps quelconque 292
2. Birapport dans un plan projectif réel 303
3. La sphère de Riemann 309
4. Birapport et configurations finies 315
A. Complément mnémotechnique sur la formule des six birapports 315
B. Géométrie projective dans une coquille de noix 318
C. Exercices 320

XI. Trois problèmes de géométrie
1. L'ellipse de Steiner 330
2. Le théorème de Desargues 333
3. L'alternative de Steiner 339
A. Exercices 353

XII. Solides platoniciens et sous-groupes finis de $\mathbf{SO(3,\mathbb{R})}$
1. Présentation 357
2. Approche topologique 359
3. Groupes d'isométries 360
4. La *toy* dualité 368
A. Annexe. Dualité des ensembles compacts convexes de $\mathbb{R}^n$ 370
B. Sous-groupes de Sylow d'un groupe d'isométries 372
C. Exercices 373

Bibliographie **377**
Index **381**

Avant-propos

Good morning, eager young minds.

A beautiful mind, Ron Howard, 2001.

L'objectif de ce livre est de raconter l'univers en termes d'orbites — ou, disons, d'actions de groupe.

Partons d'un constat simple et, pourtant, relativement peu connu : un bon nombre de théorèmes classiques d'algèbre linéaire (base incomplète et dimension, rang, Jordan, Sylvester...) sont des théorèmes de classification ; de plus, ils s'interprètent souvent comme la description des orbites pour une action particulière d'un groupe sur un ensemble.

Aussi, l'idée centrale de ce livre est l'omniprésence des groupes de transformations[1] en algèbre et en géométrie classique. Comme le note Yves André dans [3], Alexandre Grothendieck « n'hésite pas à parler de l'invention du zéro et de l'idée de groupe comme des deux plus grandes innovations mathématiques de tous les temps ».

> Il a fallu des millénaires avant que des choses aussi enfantines et omniprésentes que les groupes de symétries de certaines figures géométriques, les formes topologiques de certaines autres, le nombre zéro, les ensembles trouvent admission dans le sanctuaire [des mathématiques] !
>
> Alexandre Grothendieck, *Récoltes et semailles.*

C'est que les groupes sont la formalisation de l'idée de *symétrie*, de *transformations* qui préservent un système.

[1] À peine une ligne de texte et déjà une redondance !

Pourquoi les groupes ?

Observons les formules de changement de base en algèbre linéaire, pour un vecteur-colonne et une matrice carrée :

$$X = PX', \quad A = PA'P^{-1}.$$

On peut les interpréter de deux façons : soit comme deux objets différents reliés par une transformation, soit comme le même objet dans deux systèmes de coordonnées différents. Cette dualité des points de vue est très utile[2]. Mais, elle relève de la même situation : l'action d'un groupe (ici, le groupe linéaire) sur un ensemble (par multiplication à gauche sur les vecteurs-colonnes ou par conjugaison sur les matrices carrées).

La réduction des endomorphismes consiste, A étant donnée, à chercher P d'une part et A' aussi simple que possible d'autre part, de sorte que la relation ci-dessus soit satisfaite. En passant de A à A', on conserve certaines choses : déterminant, trace, valeurs propres... et l'on en change d'autres : les espaces propres ou caractéristiques de A sont l'image de ceux de A' par P.

Cet exemple, au cœur du programme d'algèbre de licence, est représentatif, voire emblématique, de la présence des groupes dans les problèmes de classification qui motivent le mathématicien. Ici, on « regroupe » les endomorphismes d'un espace en endomorphismes semblables ; ainsi l'ensemble des endomorphismes est partitionné par une relation d'équivalence, elle-même régie par un groupe : le groupe linéaire, qui agit par conjugaison sur l'espace des endomorphismes. Dans chaque classe d'équivalence, appelée désormais *orbite*, on sélectionne un élément représentatif, que l'on choisit « aussi simple que possible » et que l'on appellera ici *forme normale*.

Les mathématiques sont souvent affaire de classification, et classer, finalement, c'est partitionner et étiqueter. Un groupe sera vu et utilisé comme une « machine à partitionner » ; effectivement, si l'on comprend qu'une partition d'un ensemble n'est rien d'autre qu'une relation d'équivalence dans cet ensemble, on comprend alors aussi que la triade réflexif-symétrique-transitif puisse se traduire par cette autre, neutre-symétrique-associatif, qui définit la notion de groupe. L'étiquetage, de son côté, sera l'objet de la recherche d'invariants.

La notion de groupe liée à la classification et les invariants ne se rencontre pas qu'en algèbre. Pensons par exemple à une courbe paramétrée dans le plan, disons une application de classe $\mathscr{C}^1$, $\gamma : [0,1] \to \mathbb{R}^2$, $t \mapsto \big(x(t), y(t)\big)$. Avec un difféomorphisme $\varphi : [0,1] \to [0,1]$, on peut reparamétrer l'arc décrit par γ en considérant : $\tilde{\gamma} = \gamma \circ \varphi$. Mais les propriétés intéressantes,

[2] Ah ? Pourquoi ?

c'est-à-dire géométriques, sont celles qui sont invariantes par changement de paramétrage : points singuliers, longueur, rayon de courbure... Et cela constitue presque une définition de ce qu'est une propriété géométrique, à savoir : indépendante du paramétrage.

Groupes, topologie et géométrie

Pour reprendre un aphorisme de Misha Gromov, « il n'y a pas de théorème (non trivial) portant sur tous les groupes » : la structure est trop large. Aussi est-on amené à ajouter des hypothèses pour spécifier la situation et l'enrichir de théorèmes.

On commence par ajouter de la topologie : c'est une structure intermédiaire entre le « ponctuel » et le « géométrique ». L'apport de la topologie est motivé par la démonstration de théorèmes classiques (Cayley-Hamilton, par exemple) par un calcul facile valable presque partout et un argument de densité. On se place donc, dès le chapitre II, dans le large cadre des actions continues de groupes topologiques.

Mais, ce n'est guère qu'une étape vers la géométrie — au sens, ici, de la géométrie différentielle. Les méthodes qu'elle permet sont utilisées dans ce livre surtout pour remplacer quelques calculs par des concepts. Mais elles ouvrent aussi la porte à un large champ d'étude, la théorie des groupes de Lie, que l'on ne fait guère qu'effleurer.

Sur un corps fini, la topologie n'a à peu près aucun intérêt. Mais le comptage tient lieu d'étude géométrique. Cette idée apparaît déjà dans les théorèmes de Sylow : comprendre un groupe fini, cela passe par le dénombrement de ses différents sous-groupes de Sylow, voir [29]. En fait, on voit transparaître la structure géométrique observée sur les corps des complexes ou des réels dans la formule-même du cardinal de certains objets (ensemble des matrices de rang donné, espaces projectifs, grassmanniennes, hyperboloïdes...). Pour en donner l'exemple le plus simple : la formule bien connue de la série géométrique

$$\frac{q^n - 1}{q - 1} = q^{n-1} + \cdots + 1$$

pourra s'interpréter sur un corps fini $\mathbb{F}_q$ de cardinal q de la façon suivante : l'ensemble des droites vectorielles d'un espace vectoriel E, *i.e.* l'espace projectif $\mathbb{P}(E)$, est l'ensemble des orbites pour l'action libre du groupe multiplicatif $\mathbb{F}_q^\star$ sur $E\backslash\{0\}$, et il possède une décomposition en cellules. Propriété qui reste d'ailleurs valable sur $\mathbb{R}$ et $\mathbb{C}$. Jamais cette petite formule connue des bacheliers n'aura autant parlé de géométrie et de topologie !

Groupes et géométrie classique

La géométrie classique consiste à travailler sur des figures, mais qu'est ce qu'une figure ? Constatons que c'est souvent une orbite sous un groupe. Exemples : les carrés forment une orbite sous le groupe des similitudes ; les coniques non dégénérées forment une orbite sous le groupe projectif dans l'ensemble des parties du plan ; de même pour les ellipses, sous le groupe affine du plan ; pour les ellipses d'excentricité donnée, sous le groupe des similitudes ; pour les ellipses dont les longueurs d'axes sont fixées, sous le groupe des isométries planes.

Dans ce type de problèmes, il s'agira moins de classer que d'appliquer un groupe de transformations adéquat afin de se ramener, dans l'orbite des figures équivalentes, à une forme plus simple, encore une fois ce que l'on appelle sans vraiment le définir, une forme normale. C'est le principe de régularisation à l'œuvre dans l'alternative de Steiner, ou dans le théorème de Desargues, qui via l'action du groupe projectif sur la configuration de Desargues, se ramènera au théorème de Thalès, beaucoup plus simple.

Invariance

Lorsqu'il y a plusieurs orbites, il s'agit de trouver des moyens simples de les séparer, de les identifier. C'est le rôle des *invariants*. Un invariant pour une action de groupe est une application définie sur l'ensemble sur lequel le groupe opère qui est constante sur chaque orbite. Un invariant peut être un nombre (rayon d'un cercle, rang d'une application linéaire, valeurs propres...) ou un objet quelconque (genre d'une conique affine réelle, « diagramme de Young » d'une matrice nilpotente comme en III-2.3.2, « type » d'une matrice rectangulaire comme dans IV-2.3).

Un invariant est dit invariant *complet* ou *total* si l'application correspondante est injective : l'invariant prend deux valeurs différentes sur deux orbites distinctes. Par exemple, le rang est un invariant complet pour les formes quadratiques complexes, pas pour les formes quadratiques réelles (dans cette situation, le théorème de Sylvester permet de le voir et donne un invariant complet, la signature).

Bien sûr, un invariant est d'autant plus intéressant qu'il est facile à calculer et proche d'être complet. La recherche d'invariants complets est essentiellement une formulation du problème de la *classification* (des objets sur lesquels le groupe agit, à transformation — dans le groupe — près).

On peut parfois inverser le point de vue : mettre en évidence un invariant dans une situation, c'est faire naître un concept. Ainsi, un angle, c'est ce que possèdent en commun deux couples (ou paires, si l'on préfère les angles géométriques) de demi-droites (ou droites, si l'on veut) *superposables*, c'est-à-dire dans la même orbite sous le groupe des similitudes (ou des

isométries). L'ensemble des angles n'est autre que l'ensemble des orbites pour l'action du groupe des similitudes sur les couples de droites ; il devient particulièrement opérationnel lorsque l'on s'aperçoit que cet ensemble est en bijection avec $\mathbb{R}/2\pi\mathbb{Z}$, plus maniable. Le birapport constitue un exemple analogue d'invariant né de considérations de symétrie.

Un groupe, plusieurs réalisations

Certains groupes sont célèbres pour apparaître dans des contextes différents : tel est $\mathbb{R}/\mathbb{Z}$, que l'on connaît aussi comme le groupe des nombres complexes de module 1, comme celui des angles du plan, et que l'on voit aussi en étudiant l'équirépartition des suites modulo 1. C'est aussi un sous-groupe de tous les « groupes classiques ». Plus complexe, le groupe des homographies de la droite projective réelle, $\mathrm{PSL}_2(\mathbb{R})$, s'identifie aussi au groupe des isométries du plan hyperbolique (dont la droite projective est le bord) et, incidemment, au groupe des biholomorphismes complexes du disque-unité.

On verra d'autres isomorphismes de ce genre, qui permettent d'établir des dictionnaires entre géométries apparemment différentes. Par exemple, l'isomorphisme entre le groupe des homographies de la droite projective complexe et le groupe orthogonal (plus précisément la composante connexe de son neutre) d'une forme quadratique réelle de signature $(3,1)$ se comprend parfaitement si l'on considère leur action sur l'ensemble des cercles tracés sur une sphère, selon la façon de regarder les cercles (comme ensemble de points ou par leur équation).

Contenu du livre

Le livre est divisé en chapitres suffisamment indépendants entre eux pour pouvoir être lus séparément, en tout cas après avoir digéré les notions, définitions et résultats du chapitre II, autour duquel le livre est articulé. Le niveau oscille entre celui de la licence 3 et du master 2, mais la plupart des chapitres possèdent leur annexe propre où sont rappelées les notions de base (algorithme de Gauss, lemme des noyaux, décomposition de Dunford, formes quadratiques...). Seules les définitions les plus élémentaires (groupe, sous-groupe, espace vectoriel, base, morphismes...) seront supposées connues.

Pour des raisons de clarté, chaque chapitre essaie, dans la mesure du possible, de se résumer à ce qui nous paraît essentiel dans la théorie. Toutefois, on pourra trouver des compléments et développements possibles dans des sections à part, annexées au chapitre. Les développements les plus sages se trouvent dans le présent volume et les plus intrépides sont à trouver dans le second volume.

Nous avons aussi ajouté un bon nombre d'exercices à chaque chapitre : des exercices classiques de type écrit-oral d'agrégation (interne-externe), des illustrations et exemples du cours, des compléments de cours, parfois quelques notions introductives d'outils de la recherche actuelle. Dans la mesure du possible, les exercices ont été rédigés de façon que les résultats à démontrer soient clairement exprimés ; on évite au possible les questions énigmatiques qui rendent les mathématiques belles mais dévorantes, tel le Sphinx. On trouvera, pour les exercices les plus pertinents dans la préparation des examens et concours (la plupart, donc !), suffisamment d'indications pour trouver une voie directe vers la solution. Disons pour résumer que les exercices sont ici faits pour trouver des informations plutôt que pour en chercher.

Le chapitre I n'est là qu'à titre d'exemple heuristique. Il nous a paru important de démarrer sur un résultat bien connu — le théorème du rang, qui proclame que deux matrices sont équivalentes si et seulement si elles ont même rang — pour mettre en place les premières balises d'une méthode générale qui sera appliquée tout au long de l'ouvrage. En fait, il s'agira moins de découvrir ce théorème que de s'initier à une vision (le point du vue des groupes) et un vocabulaire adapté (celui des actions de groupes) qui servira par la suite. L'interprétation de ce théorème se fera donc à travers les actions de groupes, l'invariant complet (ici, le rang), la « forme normale », les algorithmes liés au calcul de l'invariant (le fameux algorithme de Gauss[3]), et enfin, sur $\mathbb{R}$ et $\mathbb{C}$, l'étude des adhérences d'orbites qui amènent à la notion d'ordre de dégénérescence (ici, les rangs sont ordonnés par l'ordre des naturels).

Le chapitre II présente *l'idée essentielle* du livre. On y trouvera les clés (définitions, vocabulaire, résultats) de tout ce qui servira par la suite, et même, n'ayons pas peur de le dire, des clés de la plus grande partie du programme universitaire (premier et deuxième cycles) d'algèbre et de géométrie. Toutes les notions étudiées *in vivo* dans le chapitre précédent seront ici définies en bonne et due forme et des résultats généraux sur les actions topologiques seront établis. Le théorème d'homéomorphisme (en deux versions, dont l'une moins générale mais facile à établir) est le théorème central du chapitre. Il donne un cadre assez général et raisonnable qui fournit un lien entre l'étude des groupes et celle de la géométrie (au sens large !), et ainsi, motive l'étude des groupes topologiques. On trouvera en annexe II-E un tableau des actions topologiques (ou pas) les plus connues (ou pas), dont le but est de synthétiser les connaissances en « algèbre-géométrie » de l'étudiant à l'agrégation (voire plus), tout en suggérant la puissance du concept. Une colonne de ce tableau est particulièrement parlante : la colonne des invariants, qui montre que la plupart des notions du programme

[3] In Gauss we trust !

peuvent se réaliser comme invariants d'actions de groupes. Le groupe et la notion d'action apparaissent comme un outil pédagogique remarquable. On constatera aussi dans ce florilège un lien fort entre la transitivité de certaines actions et des théorèmes de type « base incomplète ».

De par la familiarité (supposée) du lecteur avec la diagonalisation, le chapitre III sur la réduction paraît être un bon début pour mettre en œuvre le programme annoncé. Le problème de la réduction se résume en la détermination d'un invariant total pour l'action de conjugaison du groupe linéaire sur les endomorphismes représentés par les matrices carrées. La stratégie consiste dans un premier temps à diviser le problème en une étude d'invariants totaux, d'une part, sur les matrices diagonalisables, d'autre part, sur les matrices nilpotentes. La première étude aboutit à la notion de spectre (dans un corps algébriquement clos), en faisant toutefois attention au fait que spectre est vu comme la donnée de complexes (non nécessairement distincts) à permutation près, ce qui est sensiblement différent de la notion habituelle (mais plus naïve) d'« ensemble de valeurs propres »[4]. La seconde étude aboutit à la notion de diagramme de Young comme invariant, et les formes de Jordan y font office de formes normales. C'est la décomposition de Dunford qui va permettre de fournir une réponse au problème global (sur les endomorphismes) à partir des deux solutions partielles (à partir des endomorphismes diagonalisables et nilpotents). La compréhension des adhérences d'orbites se fait aussi en deux temps : le cas diagonalisable ne fournit que des orbites fermées, lesquelles ne posent donc aucun problème ; en revanche, le cas nilpotent débouche sur l'ordre de Chevalley pour lequel le choix des diagrammes de Young est particulièrement pertinent. Nous donnons ensuite une version valable sur tous les corps avec les invariants de similitude. On en profitera pour parler des notions de semi-simplicité, qui généralise la diagonalisabilité. La notion de commutant, liée aux stabilisateurs d'orbites, sera étudiée dans le second volume.

Le chapitre IV qui suit part d'un problème omniprésent en algèbre linéaire : celui de la résolution des systèmes linéaires. La méthode dite du pivot de Gauss s'interprète alors comme une étude de formes normales pour l'action du groupe linéaire (encore lui) par multiplication à gauche sur l'espace des matrices. Ici, la recherche d'un invariant aboutit à la notion de noyau et, de façon plus concrète et combinatoire, à la notion de matrice échelonnée réduite. Une version duale débouche sur la notion d'image et de matrice co-échelonnée. On a ensuite besoin de comprendre l'espace topologique dans lequel ces invariants évoluent, et l'on étudie alors la grassmannienne ; noyaux et images sont des sous-espaces, et il devient maintenant naturel d'étudier leur ensemble comme espace topologique. On présente l'outil permettant de trivialiser localement sa topologie : la décomposition en cellules.

[4]Certains auteurs en font la distinction en parlant de *spectre* et de *spectre complet*.

Le chapitre V est le dernier du présent volume traitant de classification. Il s'agit là de la classification des formes bilinéaires et plus précisément des formes bilinéaires symétriques, qui permettent d'étudier les formes quadratiques. Encore une fois, on se ramène à l'étude de l'action du groupe linéaire, mais cette fois-ci par congruence sur les matrices symétriques. On se rend compte que l'invariant total dépend drastiquement du corps de base choisi. On se contentera de fournir de tels invariants pour les corps, $\mathbb{C}$, $\mathbb{R}$ et $\mathbb{F}_q$. On retrouve alors des invariants, comme le rang, la signature, et le discriminant. Une fois les orbites classées, on s'intéresse à leurs stabilisateurs, qui vont fournir les derniers « groupes classiques » qui nous poursuivront tout au long de l'ouvrage : les groupes orthogonaux. Nous en profiterons pour étudier leur action naturelle sur l'espace vectoriel ambiant. On verra qu'il n'y a pas de « théorème de base incomplète » adapté au contexte, et que la transitivité des actions dépend d'un théorème plus crochu : le théorème de Witt.

Les chapitres qui suivent ont pour but de développer des outils permettant d'étudier les groupes classiques pour eux-mêmes, selon le principe du « reculer pour mieux sauter ». Le chapitre VI introduit la décomposition polaire qui peut être vu comme un théorème de séparation topologique. Utiliser la décomposition polaire sur un groupe, c'est diviser son étude topologique en deux études plus simples, celle d'un sous-groupe compact maximal et celle d'un espace topologique homéomorphe à un espace vectoriel. Nous donnerons en développement et en exercices plusieurs études de groupes utilisant cette décomposition. L'étude de l'exponentielle trouvera tout à fait sa place dans ce chapitre.

Puis, le chapitre VII fournira un corps non commutatif, le corps des quaternions, permettant de comprendre les groupes $\mathrm{SO}(3)$ et $\mathrm{SO}(4)$ à partir de la sphère $\mathbf{S}^3$ de dimension 3 des quaternions, elle-même munie d'une structure de groupe dans lequel le calcul est plus aisé.

On illustre ensuite dans le chapitre VIII comment combinatoire et actions de groupes font bon ménage lorsque l'on travaille sur les corps finis. Dans ce contexte, le théorème d'homéomorphisme du chapitre II est réduit à sa plus simple expression : le lemme du berger. Malgré la simplicité des méthodes utilisées, on pourra noter la remarquable efficacité des actions de groupes pour le dénombrement d'objets géométriques. De plus, par des arguments de comptage (injectif et même cardinal impliquent ensemble bijectif), on obtiendra des isomorphismes, dits exceptionnels, entre certains groupes classiques, qui établissent des correspondances utiles entre des domaines *a priori* différents. Ce chapitre élémentaire, qui aurait pu tout aussi bien être placé au tout début de l'ouvrage, sert ici de préliminaire au chapitre qui suit.

Le chapitre IX sur les groupes de Lie suit la même problématique, mais à présent le corps de base est $\mathbb{R}$ ou $\mathbb{C}$: du coup, les arguments de dénombrement n'ont plus cours pour montrer la surjectivité par une méthode générale. On constatera alors la performance du calcul différentiel joint à la théorie des groupes : on ajoute une structure de variété différentielle compatible avec la loi de groupe et le théorème d'inversion locale permet d'obtenir cette surjectivité si souvent convoitée en mathématiques. Ici, la recherche d'un isomorphisme exceptionnel est surtout prétexte à l'introduction des méthodes différentielles. Néanmoins, on peut interpréter un tel isomorphisme comme l'incarnation d'un dictionnaire entre deux géométries.

Les trois derniers chapitres relèvent d'une géométrie plus traditionnelle, disons au moins dans la problématique. Le chapitre X présente en détails la droite projective $\mathbb{P}^1(\mathbb{K})$ (où $\mathbb{K}$ désigne un corps quelconque), que l'on peut considérer comme un espace gouverné par le groupe projectif $\mathrm{PGL}_2(\mathbb{K})$. Un triplet de points deux à deux distincts constitue ce qu'on appelle un repère projectif. Cette définition s'inscrit dans une philosophie générale qui s'exprime ici par le fait que le groupe projectif agit de façon simplement transitive sur l'ensemble de ces triplets. Ainsi, l'action cesse d'être transitive si l'on fait agir le groupe projectif sur quatre points. Idée simple et féconde qui fournira, puisqu'il y a désormais plusieurs orbites, un invariant intéressant : le birapport de quatre points. Ce birapport est ensuite étudié sous plusieurs angles (sic !).

On résout dans le chapitre XI trois problèmes de géométrie plane. Il s'agit là de donner trois exemples de groupes de transformations adaptés qui permettent de ramener un problème géométrique donné à une forme normale, c'est-à-dire particulièrement agréable à étudier. On expliquera en termes d'invariants pourquoi choisir ici le groupe affine, là, le groupe projectif réel, et enfin le groupe projectif complexe.

Pour finir, le dernier chapitre s'intéresse aux célèbres solides platoniciens. Assez différent des autres chapitres, il a tout de même sa place ici en raison des liens étroits qu'il présente entre groupes et géométrie. On calculera les groupes d'isométries de ces solides et l'on verra comment un simple coup d'œil sur ces objets permet d'obtenir de jolis morphismes, voire isomorphismes entre groupes classiques. Mais ceci n'est pas une fin, puisque l'on reviendra longuement, profondément, voire largement sur le sujet dans le second tome, avec comme outil la théorie des représentations...

Références de base

Ce livre a été écrit sous l'influence majeure de Daniel Perrin, Michèle Audin et Rached Mneimné, et en particulier de leur ouvrages respectifs [29], [4] et [25], dont la lecture est plus que chaudement recommandée. Qu'ils nous pardonnent de nous abriter derrière eux. Les deux auteurs ont été fortement

influencés par la magistrale présentation de la géométrie dans l'esprit du programme d'Erlangen que l'on trouve dans la première annexe de [18].

L'expérience de l'enseignement a été une autre source d'inspiration. C'est durant la période des mouvements universitaires de 2008-2009 contre la loi LRU que les étudiants de master de l'université Lyon 1 ont eu l'idée de se réunir sur un forum internet pour travailler en s'entraidant, épaulés également par les enseignant-chercheurs, dans une ambiance sympathique et parfois bien nocturne. Ce forum a été un outil majeur, qui nous a permis de voir quels points du programme d'agrégation posaient des problèmes aux étudiants. Un premier polycopié est né de cette expérience, et ce dernier n'a cessé de s'enrichir d'année en année grâce aux remarques des agrégatifs qui l'utilisaient.

Remerciements

Nous remercions vivement Serge Parmentier et Kenji Iohara, qui nous ont fourni un grand nombre d'exercices ; mention spéciale à Johann Colombano-Rut, dont les notes de cours en LaTeX avec de belles illustrations ont fourni une première version de ce qui est devenu ce livre ; à Florent Hours pour son investissement considérable dans le forum des étudiants. Nous remercions également nos prédécesseurs : les professeurs qui ont enseigné le cours de « Groupes et géométrie » durant les années précédentes : Claude Roger, Gilbert Hector, et notre regretté collègue Fokko Du Cloux. Nous tenons à remercier les étudiants qui, soit par des questions insistantes et bien fondées, soit par leurs remarques avisées, voire une lecture minutieuse, ont permis d'améliorer le livre : Benjamin Bertrand, Jérémy Legendre, François Viard, Hélène Mariette, Maxime Moine (du Haut-Bugey), Teddy Mignot, Julie Carlier, et les autres. Un grand merci à Bruno Calado pour sa relecture des premiers chapitres et leurs annexes, il a su fureter dans les arcanes des preuves topologiques et a découvert des erreurs inavouables. Merci aussi à Amaury Thuillier, pour plusieurs conversations et suggestions utiles et enfin à Michel Cretin, préparateur à l'agrégation de Lyon 1 pour son intérêt dans notre projet. Merci, enfin, à Alberto Arabia et aux éditions Calvage & Mounet pour l'accueil et l'attention toute particulière qu'ils ont réservés à cet ouvrage.

Signalisation

Nous avons placé ça et là des signalisations qu'il est bon de savoir décoder. Voici une liste non exhaustive.

Danger classique.

Attention à rester dans le cadre des hypothèses. Une généralisation abusive pourrait être fatale.

On dit qu'un sentiment de honte est vite passé, mais c'était avant Youtube. Évitez les peaux de banane !

Concept délicat, à manipuler avec précaution si l'on ne veut pas se retrouver au fond du ravin.

Concept plus métaphysique, qui peut engendrer des désordres existentiels.

Pourquoi ne pas faire une petite pause et admirer le paysage ?

Résultat de bon niveau, qui peut s'exposer avec brio lors d'un oral de concours.

Alternative.

Résultat bien sympathique, mais dont la preuve nous passera au-dessus de la tête. Pour ceux qui aiment prendre de l'altitude.

Avertissement : ici et là, les notations et les concepts pourront subir quelques fluctuations un peu erratiques. Qu'on le pardonne aux auteurs en se rappelant la phrase d'Oscar Wilde : *Consistency is the last refuge of the unimaginative.*

Chapitre I

Actions et théorèmes du rang

Everything begins with choice.

Matrix Reloaded, Andy & Larry Wachowski, 2003.

Le but de ce chapitre est de motiver l'étude des actions de groupes, qui seront introduites au chapitre suivant. Le théorème du rang apparaît dès les premiers cours d'algèbre linéaire : on le déduit en général du théorème de la base incomplète et de l'invariance du cardinal d'une base d'un espace. Une de ses formulations affirme que deux matrices sont équivalentes si et seulement si elles ont le même rang. On revisite ce théorème fameux en termes d'invariants pour l'action d'un groupe particulier.

Le théorème du rang est valable sur n'importe quel corps. Cependant, comme dans le reste du livre, on va se restreindre rapidement aux exemples suivants : les corps finis, ce qui donne lieu à des dénombrements non triviaux ; les corps des réels et des complexes, ce qui permet de parler facilement[1] de topologie. C'est, de plus, l'occasion de motiver et de voir dans cet exemple précis certains ingrédients d'une action de groupe : orbite, stabilisateur...

[1] Ce n'est pas la seule façon. Pour presque toutes les situations étudiées dans ce livre, on pourrait aussi utiliser la topologie de Zariski et, ainsi, travailler sur n'importe quel corps. *A contrario*, on pourrait s'intéresser à des corps munis de topologies compliquées, comme le corps des nombres p-adiques.

1. Théorème du rang

Soit $\mathbb{K}$ un corps. Pour tout entier naturel n, on munit $\mathbb{K}^n$ de sa structure naturelle[2] d'espace vectoriel. Soit m, n deux entiers naturels non nuls. Soit $\varphi : \mathbb{K}^n \to \mathbb{K}^m$ une application linéaire. Choisissons deux bases $\mathbf{e} = (e_j)_{1\leqslant j\leqslant n}$ et $\mathbf{f} = (f_i)_{1\leqslant i\leqslant m}$ de $\mathbb{K}^n$ et $\mathbb{K}^m$ respectivement. Considérons l'espace vectoriel $\mathscr{M}_{m,n}(\mathbb{K})$ des matrices à m lignes et n colonnes, dites matrices de taille $m \times n$, à coefficients dans $\mathbb{K}$, et soit $A = \mathrm{mat}_{\mathbf{e},\mathbf{f}}(\varphi) \in \mathscr{M}_{m,n}(\mathbb{K})$ la matrice de φ dans les bases $\mathbf{e}$ et $\mathbf{f}$, c'est-à-dire la matrice dont les colonnes sont données par les vecteurs $\varphi(e_j)$ dans la base $\mathbf{f}$.

Il est plaisant d'associer un outil pratique (la matrice, sur laquelle on peut faire des calculs) à un objet abstrait (l'application linéaire), mais cette correspondance soulève le problème du *choix des bases*. Or, on sait qu'en mathématiques, il n'est pas de choix anodin.

En l'espèce, rappelons la formule de changement de bases : si l'on choisit deux autres bases $\mathbf{e}'$ et $\mathbf{f}'$ de $\mathbb{K}^n$ et $\mathbb{K}^m$, si l'on note P et Q les matrices de passage de $\mathbf{e}$ à $\mathbf{e}'$ et de $\mathbf{f}$ à $\mathbf{f}'$, la matrice de φ dans les bases $\mathbf{e}$ et $\mathbf{f}$ est : $\mathrm{mat}_{\mathbf{e}',\mathbf{f}'}(\varphi) = Q^{-1}AP$.

Cela incite à introduire une relation d'équivalence appelée pertinemment (ou pas !) *équivalence*.

1.1. Définition. On dit que deux matrices A et B de même format sont *équivalentes* si elles représentent la même application linéaire ou, ce qui revient au même, s'il existe deux matrices inversibles P et Q telles que $A = Q^{-1}BP$. On note alors parfois : $A \approx B$.

1.2. Rappel (rang). Avec les notations ci-dessus, on définit le rang d'une matrice A comme la dimension de l'espace $\mathrm{col}(A)$ engendré par les colonnes de A dans $\mathbb{K}^m$ et le rang de φ comme la dimension de l'image de φ. On montre que ces entiers sont égaux :

$$\mathrm{rg}\, A = \dim \mathrm{col}(A) = \dim \mathrm{Im}\, \varphi = \mathrm{rg}\, \varphi.$$

Le théorème suivant donne un critère pour que deux matrices soient équivalentes. Il peut se résumer en disant que le rang est un invariant total pour l'action de $\mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$ sur $\mathscr{M}_{m,n}(\mathbb{K})$ qui sera définie au paragraphe 2. Comme un certain nombre de résultats dits de transitivité qui seront étudiés dans les chapitres suivants, la preuve fait intervenir le *théorème de la base incomplète*, qui dit que dans un espace vectoriel de dimension finie, toute partie libre peut être complétée en une base.

[2] Pour $n = 0$, l'espace $\mathbb{K}^0$ est le singleton $\{0\}$, où 0 est l'application de l'ensemble vide vers $\mathbb{K}$. Il y a donc bien une structure d'espace vectoriel sur $\mathbb{K}^0$.

1.3. Théorème (du rang). *Soit $\mathbb{K}$ un corps, m et n deux entiers, A et B deux matrices $m \times n$ à coefficients dans $\mathbb{K}$. Alors, A et B sont équivalentes si et seulement si elles ont le même rang :* $\operatorname{rg} A = \operatorname{rg} B$.

Démonstration. Supposons que A et B sont équivalentes, c'est-à-dire qu'elles représentent la même application linéaire φ. Alors, les rangs de A et B coïncident avec le rang de φ, donc ils sont égaux.

Réciproquement, si A est une matrice de rang r, on considère $\mathbf{e}$, $\mathbf{f}$ deux bases et φ une application linéaire[3] telles que $A = \mathrm{mat}_{\mathbf{e},\mathbf{f}}(\varphi)$. Soit I_r la matrice identité $r \times r$ et $\mathrm{I}_{m,n,r}$ la matrice $m \times n$ décrite par blocs de la façon suivante :

$$\mathrm{I}_{m,n,r} = \begin{pmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Il suffit de montrer que A est équivalente à $\mathrm{I}_{m,n,r}$. Alors, de même, B sera aussi équivalente à $\mathrm{I}_{m,n,r}$ et, par symétrie et transitivité, on en déduira que les matrices A et B sont équivalentes.

Soit $(e''_1, \ldots, e''_s)$ une base du noyau de φ. Par le théorème de la base incomplète, on complète cette famille libre en une base de $\mathbb{K}^n$, notée $\mathbf{e}' = (e'_1, \ldots, e'_t, e''_1, \ldots, e''_s)$. Notons E' l'espace engendré par $e'_1, \ldots, e'_t$ (l'entier t sera ici le rang de φ). De plus, pour i entier compris entre 1 et t, notons f'_i l'image de e'_i par φ. La restriction de φ à E' est injective, car son noyau est l'intersection, triviale, de E' et du noyau de φ. Comme, de plus, la famille $(e'_i)_{1 \leqslant i \leqslant t}$ est libre, la famille $(f'_i)_{1 \leqslant i \leqslant t}$ est libre. On la complète en une base $\mathbf{f}' = (f'_i)_{1 \leqslant i \leqslant m}$ de $\mathbb{K}^m$. On a alors : $\mathrm{mat}_{\mathbf{e}',\mathbf{f}'}(\varphi) = \mathrm{I}_{m,n,r}$, si bien que A et $\mathrm{I}_{m,n,r}$ sont équivalentes, comme annoncé. □

1.4. Remarque. De cette démonstration, on déduit de plus les égalités : $s + t = n$ et $t = \operatorname{rg} \varphi$, avec $s = \dim \operatorname{Ker} \varphi$. C'est la formulation la plus classique du théorème du rang.

2. Action de $\mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$ sur $\mathscr{M}_{m,n}(\mathbb{K})$ par équivalence

Le problème de l'arbitraire du choix des bases que nous venons de soulever se décrit par l'action particulière d'un groupe. Le vocabulaire général des actions de groupes est présenté en annexe A.

En mathématiques, les exemples abondent là où un problème de choix implique la considération d'une « orbite des possibles ». Pour n'en citer qu'un,

[3] On peut par exemple prendre pour $\mathbf{e}$ (resp. $\mathbf{f}$) la base canonique de $\mathbb{K}^n$ (resp. $\mathbb{K}^m$) et $\varphi = \varphi_A : \mathbb{K}^n \to \mathbb{K}^m$, $X \mapsto AX$.

en théorie de Galois[4], le choix d'une solution d'une équation polynomiale n'est pas naturel. Mais le groupe de Galois est un objet naturel qui rend compte, en permutant les solutions, de la multiplicité des choix possibles.

2.1. Définition. Pour $\mathbb{K}$ un corps et d entier naturel non nul, on appelle *groupe linéaire* d'ordre d et l'on note $\mathrm{GL}_d(\mathbb{K})$ le groupe des matrices inversibles de taille $d \times d$.

Dans la suite, on notera G le produit direct de groupes $\mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$. Considérons l'application :

$$\begin{aligned} \alpha : G \times \mathscr{M}_{m,n}(\mathbb{K}) &\longrightarrow \mathscr{M}_{m,n}(\mathbb{K}) \\ \big((P,Q),A\big) &\longmapsto (P,Q)\cdot A = PAQ^{-1}. \end{aligned}$$

Pour P et P' de $\mathrm{GL}_m(\mathbb{K})$, pour Q et Q' de $\mathrm{GL}_n(\mathbb{K})$ et pour A de $\mathscr{M}_{m,n}(\mathbb{K})$, on a une relation pouvant évoquer l'associativité :

$$[(P,Q)(P',Q')]\cdot A = PP'A(QQ')^{-1} = (P,Q)\cdot[(P',Q')\cdot A]\,.$$

En particulier, si I_m et I_n sont les matrices identités : $(\mathrm{I}_m,\mathrm{I}_n)\cdot A = A$.

L'application α définit ce que l'on appelle une action à gauche, ce qui justifie la notation. De même, on obtiendrait clairement une action à droite en posant : $A\cdot(P,Q) = P^{-1}AQ$.

La reformulation du théorème du rang passe par l'identification de la relation d'équivalence $\approx$ avec la relation « être dans la même orbite sous l'action de G ». Explicitons ce fait.

2.2. Définition. On définit l'*orbite* d'une matrice A sous l'action de G comme l'ensemble

$$G\cdot A = \{(P,Q)\cdot A \,:\, (P,Q)\in G\}.$$

Considérons l'orbite d'une matrice A sous l'action de G :

$$\begin{aligned} G\cdot A &= \{B\in\mathscr{M}_{m,n}(\mathbb{K}) \,:\, \exists (P,Q)\in G,\ B = PAQ^{-1}\} \\ &= \{B\in\mathscr{M}_{m,n}(\mathbb{K}) \,:\, B \text{ est équivalente à } A\} \\ &= \{B\in\mathscr{M}_{m,n}(\mathbb{K}) \,:\, \operatorname{rg} B = \operatorname{rg} A\}. \end{aligned}$$

Étant donné une application linéaire φ, plutôt que lui associer *une* matrice, il est donc plus naturel de lui associer *toutes* les matrices qui représentent φ dans des bases convenables, c'est-à-dire une orbite sous l'action de G, ou

[4]La théorie de Galois est une « théorie de l'ambiguïté », pour reprendre l'expression de Galois lui-même ([3]).

encore une classe d'équivalence pour la relation $\approx$. Ici, le mot *naturel* signifie « indépendant de tout choix ».

On peut reformuler le théorème du rang.

2.3. Théorème (du rang, autre formulation). *Deux matrices A et B de $\mathscr{M}_{m,n}(\mathbb{K})$ sont dans la même orbite sous l'action de $G = \mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$ si et seulement si elles ont le même rang :*

$$G \cdot A = G \cdot B \iff \mathrm{rg}\, A = \mathrm{rg}\, B.$$

En particulier, les orbites sont paramétrées par le rang, entier compris entre 0 et $\min(m,n)$.

3. Sous-groupe stabilisateur, bijection entre quotient et orbite

Dans cette section, nous établissons un lien fondamental entre groupes et géométries, c'est-à-dire entre un groupe G et un objet sur lequel G agit transitivement. Ce lien s'exprimera sous la forme d'une bijection entre un quotient G/H du groupe et une orbite de G.

Pour motiver cette partie, nous continuons l'étude de l'action de $G = \mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$ sur $\mathscr{M}_{m,n}(\mathbb{K})$ par équivalence, mais en supposant le corps $\mathbb{K}$ fini. Nous voulons calculer le cardinal d'une orbite, c'est-à-dire ici, le cardinal de l'ensemble des matrices de taille $m \times n$ et de rang r.

Toutefois, la plupart des résultats énoncés (tel le théorème 3.2) valent pour toute action de groupe. Fixons A dans $\mathscr{M}_{m,n}(\mathbb{K})$. Considérons l'application suivante, qui est surjective par définition :

$$\begin{aligned} \alpha_A : G &\longrightarrow G \cdot A \\ g &\longmapsto g \cdot A. \end{aligned}$$

L'application α_A est-elle injective ? Soit g et g' deux éléments de G, on a :

$$\begin{aligned} \alpha_A(g) = \alpha_A(g') &\iff g \cdot A = g' \cdot A \\ &\iff g^{-1} \cdot (g \cdot A) = g^{-1} \cdot (g' \cdot A) \\ &\iff (g^{-1}g) \cdot A = (g^{-1}g') \cdot A \\ &\iff (g^{-1}g') \cdot A = A. \end{aligned}$$

On voit ainsi que l'injectivité de α_A est liée à l'existence de stabilisateurs, au sens suivant qui sera précisé en annexe dans un cadre général...

3.1. Définition. On appelle *stabilisateur* de A et l'on note $\mathrm{Stab}_G(A)$ ou encore G_A l'ensemble :

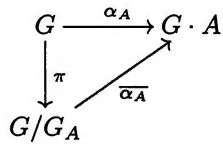
$$\begin{aligned}\mathrm{Stab}_G(A) = G_A &= \{h \in G\ ,\quad h \cdot A = A\} \\ &= \{(P,Q) \in \mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})\ ,\quad PAQ^{-1} = A\}.\end{aligned}$$

On vérifie que le stabilisateur d'un point est un sous-groupe de G. En général, il n'est pas distingué (voir la proposition 3.5, ci-dessous).

Le calcul ci-dessus montre que g et g' ont la même image par α_A si et seulement si ils appartiennent à la même classe à gauche modulo G_A, c'est-à-dire si les images de g et g' dans G/G_A coïncident.

3.2. Théorème. *Soit G un groupe agissant sur un ensemble X et A un élément de X. On note G_A le stabilisateur de A, et $\pi : G \to G/G_A$ la projection canonique et α_A l'application $G \to G \cdot A$, $g \mapsto g \cdot A$.*

Alors, il existe une unique application $\overline{\alpha_A} : G/G_A \to G \cdot A$ qui rend commutatif le diagramme :



L'application $\overline{\alpha_A}$ est bijective.

Démonstration. Puisque π est surjective, si $\overline{\alpha_A}$ existe, elle est unique. En effet, fixons un élément γ de G/G_A. On choisit un antécédent g de γ par l'application π. Alors :

$$(*) \qquad \overline{\alpha_A}(\gamma) = \alpha_A(g).$$

Cet élément de $G \cdot A$ est uniquement déterminé par α_A.

D'après ce qui précède, pour construire $\overline{\alpha_A}$, il suffit de montrer que $\alpha_A(g)$ ne dépend que de γ, et pas du choix de g. Soit g' un autre antécédent de γ par π : par définition de π, $gG_A = g'G_A$, ce qui signifie qu'il existe h dans G_A tel que $g' = gh$. Mais alors, puisque $h \cdot A = A$, il vient :

$$\alpha_A(g') = g' \cdot A = (gh) \cdot A = g \cdot (h \cdot A) = g \cdot A = \alpha_A(g).$$

Cela a donc bien un sens d'utiliser $(*)$ pour définir $\overline{\alpha}_A(\gamma)$. L'existence de $\overline{\alpha}_A$ en résulte. De plus, elle est surjective par construction et injective par l'étude qui précède. □

3.3. Mise en garde (valable pour toute action de groupe). Le théorème met en évidence une bijection du quotient $\big(\mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})\big)/G_A$ vers l'orbite $G \cdot A$ mais ce n'est pas un isomorphisme de groupes. Et pour cause, en général, le stabilisateur G_A n'est pas distingué (voir la proposition 3.5) et la partie $G \cdot A$ de l'ensemble X n'a aucune raison d'être un groupe non plus.

Dans le cas particulier de l'action de $\mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$ sur $\mathscr{M}_{m,n}(\mathbb{K})$, on en déduit le résultat suivant.

3.4. Corollaire. *Soit $\mathbb{K}$ un corps fini, m et n deux entiers, $G = \mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$ et $A \in \mathscr{M}_{m,n}(\mathbb{K})$. Alors :*

$$|G \cdot A| = \frac{|\mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})|}{|G_A|}.$$

Pour calculer le cardinal d'une orbite dans le cas d'un corps fini, il serait donc utile de déterminer le (cardinal du) stabilisateur. « Le » stabilisateur ? De nouveau, nous semblons avoir un problème de choix d'un élément dans l'orbite : espérons que le stabilisateur ne dépende « pas trop » du point. Du moins, que son cardinal ne dépende pas du point dans l'orbite. Nous tombons ici sur la première manifestation du très général « principe de conjugaison », cher à Daniel Perrin ([29]) et Michèle Audin ([4]).

3.5. Proposition. *Soit G un groupe agissant sur un ensemble X, A un élément de X, g un élément de G, G_A et $G_{g\cdot A}$ les stabilisateurs de A et $g\cdot A$. Alors :*

$$G_{g\cdot A} = gG_Ag^{-1}.$$

En particulier, les stabilisateurs de deux éléments de la même orbite sont des groupes isomorphes.

Démonstration. Soit h un élément de G. On a :

$$\begin{aligned} h \in G_{g\cdot A} &\iff h \cdot (g \cdot A) = g \cdot A \\ &\iff hg \cdot A = g \cdot A \\ &\iff g^{-1}hg \cdot A = A \\ &\iff g^{-1}hg \in G_A \\ &\iff h \in gG_Ag^{-1}. \end{aligned}$$

Ainsi, on a : $G_{g\cdot A} = gG_Ag^{-1}$ et la conjugaison $G_A \to G_{g\cdot A}$, $h \mapsto ghg^{-1}$ est un isomorphisme. □

Revenons à l'action de $\mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$ sur $\mathscr{M}_{m,n}(\mathbb{K})$ lorsque $\mathbb{K}$ est un corps fini. On vient de constater que pour calculer le cardinal de l'orbite des matrices de rang r fixé, on peut choisir n'importe quel élément de l'orbite et calculer le cardinal de son stabilisateur. Il est naturel de choisir un élément « aussi simple que possible », dans notre cas la matrice $\mathrm{I}_{m,n,r}$.

Déterminons le stabilisateur de $\mathrm{I}_{m,n,r}$. Les matrices ci-dessous sont découpées en blocs de sorte que le bloc supérieur gauche soit de taille $r \times r$. Soit $(P, Q) \in G$. On écrit :

$$P = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad Q = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix}.$$

Pour tester si (P, Q) stabilise $\mathrm{I}_{m,n,r}$, il est astucieux de ne pas calculer l'inverse de Q, ce qui serait difficile :

$$\begin{aligned} P\mathrm{I}_{m,n,r}Q^{-1} = \mathrm{I}_{m,n,r} &\iff P\mathrm{I}_{m,n,r} = \mathrm{I}_{m,n,r}Q \\ &\iff \begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} \\ &\iff \begin{pmatrix} A & 0 \\ C & 0 \end{pmatrix} = \begin{pmatrix} A' & B' \\ 0 & 0 \end{pmatrix} \\ &\iff A = A' \text{ et } C = 0 \text{ et } B' = 0 \\ &\iff P = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \text{ et } Q = \begin{pmatrix} A & 0 \\ C' & D' \end{pmatrix}. \end{aligned}$$

Puisque le déterminant de P (resp. Q) est ici le produit des déterminants des blocs diagonaux, P (resp. Q) est inversible si et seulement si A et D (resp. A et D') le sont. D'autre part, B (resp. C') est quelconque dans $\mathscr{M}_{r,m-r}(\mathbb{K})$ (resp. $\mathscr{M}_{n-r,r}(\mathbb{K})$).

On saura, voir exercice II-5.3.6, comment décrire le stabilisateur $G_{\mathrm{I}_{m,n,r}}$ comme un produit semi-direct :

$$(\mathscr{M}_{r,m-r}(\mathbb{K}) \times \mathscr{M}_{n-r,r}(\mathbb{K})) \rtimes \big(\mathrm{GL}_r(\mathbb{K}) \times \mathrm{GL}_{m-r}(\mathbb{K}) \times \mathrm{GL}_{n-r}(\mathbb{K})\big),$$

sous-groupe du groupe décrit par :

$$\begin{pmatrix} \mathrm{GL}_r(\mathbb{K}) & \mathscr{M}_{r,m-r}(\mathbb{K}) \\ 0 & \mathrm{GL}_{m-r}(\mathbb{K}) \end{pmatrix} \times \begin{pmatrix} \mathrm{GL}_r(\mathbb{K}) & 0 \\ \mathscr{M}_{n-r,r}(\mathbb{K}) & \mathrm{GL}_{n-r}(\mathbb{K}) \end{pmatrix}.$$

3.6. Exercice. *Pour d entier naturel et $\mathbb{K}$ corps fini de cardinal q. Montrer que l'étude ci-dessus permet de calculer les cardinaux des orbites à partir*

du cardinal du groupe. Plus précisément, montrer la relation :

$$|\operatorname{GL}_d(\mathbb{K})| = \prod_{i=0}^{d-1}(q^d - q^i),$$

par un dénombrement de l'ensemble des bases de $\mathbb{K}^d$, et que le nombre de matrices de taille $m \times n$ de rang r est

$$|G \cdot \mathrm{I}_{m,n,r}| = q^{a_{m,n,r}} \prod_{i=1}^{m-r}(q^{r+i} - 1) \prod_{i=1}^{n-r}(q^{r+i} - 1),$$

où $a_{m,n,r} = \frac{1}{2}\big((m-r)(m-r-1) + (n-r)(n-r-1)\big)$.

On remarquera en passant la symétrie en m et n de la formule, justifiée par le fait que la transposition conserve le rang.

4. Topologie des orbites (𝕂 corps des réels ou des complexes)

Lorsque $\mathbb{K}$ est $\mathbb{R}$ ou $\mathbb{C}$, tous les espaces de matrices sont munis de la topologie d'espace vectoriel normé de dimension finie : il y a beaucoup de normes, mais elles induisent toutes la même topologie. La somme et le produit de matrices ne font intervenir que des polynômes en les coefficients des matrices, elles sont donc continues pour cette topologie.

Pour les lecteurs à qui l'introduction de la topologie dans un contexte d'algèbre linéaire semble artificielle, on ne peut guère que les exhorter à la patience : plus loin, inversant le point de vue, nous montrerons que la topologie permet de montrer élégamment de puissants théorèmes.

Pour l'heure, jugeons qu'il est raisonnablement naturel d'étudier la topologie des orbites dans $\mathscr{M}_{m,n}(\mathbb{K})$ sous l'action de $\operatorname{GL}_m(\mathbb{K}) \times \operatorname{GL}_n(\mathbb{K})$. En général, elles ne sont ni ouvertes, ni fermées, mais l'on peut en déterminer l'adhérence.

4.1. Proposition. *Soit $\mathbb{K}$ le corps des réels ou des complexes. Soit m et n deux entiers. Pour r entiers satisfaisant à $0 \leqslant r \leqslant \min(m,n)$, on note $\mathscr{O}_r$ l'orbite des matrices $m \times n$ de rang r à coefficients dans $\mathbb{K}$. Alors, l'adhérence $\overline{\mathscr{O}_r}$ de l'orbite est donnée par :*

$$\overline{\mathscr{O}_r} = \bigcup_{0 \leqslant k \leqslant r} \mathscr{O}_k \quad \textit{(réunion disjointe).}$$

4.2. Remarque. Ce qui s'appelle mettre de l'ordre dans les rangs... Blague à part, la proposition fournit le premier *ordre de dégénérescence* présenté dans l'ouvrage : un ordre sur les valeurs de l'invariant total (ici les rangs possibles), ou plus directement sur l'ensemble des orbites, induit par les adhérences d'orbites. L'ordre de dégénérescence, vu sur l'ensemble des orbites, est défini ainsi : l'orbite $\mathscr{O}$ sera plus petite qu'une orbite $\mathscr{O}'$ si $\mathscr{O}$ est incluse dans l'adhérence de $\mathscr{O}'$. On en verra divers exemple dans la suite (en particulier dans le chapitre IV avec les adhérences des B-orbites dans la grassmannienne et dans le chapitre III avec l'ordre de Chevalley sur les matrices nilpotentes). Attention ! Ce que n'aura pas manqué le lecteur assidu, c'est qu'il n'est pas clair du tout que la relation ainsi définie sur l'ensemble des orbites soit toujours une relation d'ordre. Mais ce qui fait marcher la chose, c'est que dans chaque exemple étudié, les orbites sont localement fermées, *i.e.* elles sont ouvertes dans leur adhérence.

Pour la preuve, introduisons quelques notations relatives aux mineurs. Étant donné deux parties $I \subset \{1, \dots, m\}$ et $J \subset \{1, \dots, n\}$ de *même cardinal*, le mineur d'indice (I, J) est l'application

$$\Delta_{I,J} : \mathscr{M}_{m,n}(\mathbb{K}) \longrightarrow \mathbb{K}, \quad (a_{i,j})_{1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n} \longmapsto \det(a_{i,j})_{i \in I, j \in J}.$$

Un théorème important d'algèbre linéaire caractérise le rang d'une matrice A comme l'ordre du plus grand mineur non nul :

$$\operatorname{rg} A = \max\{r \in \mathbb{N}\ ,\ \exists I,\ \exists J,\ |I| = |J| = r \text{ et } \Delta_{I,J}(A) \neq 0\}.$$

Démonstration. Montrons d'abord que la réunion des $\mathscr{O}_k$ $(0 \leqslant k \leqslant r)$ est un fermé de $\mathscr{M}_{m,n}(\mathbb{K})$. Notons-la provisoirement $\widehat{\mathscr{O}}_r$. D'après ce qui précède, le rang d'une matrice A est au plus r si et seulement si tous ses mineurs d'ordre supérieur ou égaux à $r+1$ sont nuls. (En fait, en utilisant le développement du déterminant par rapport à une colonne, on voit que les mineurs d'ordre $r+1$ suffisent.) En symboles :

$$\widehat{\mathscr{O}}_r = \bigcap_{|I|=|J| \geqslant r+1} \Delta_{I,J}^{-1}(0).$$

Par continuité des fonctions $\Delta_{I,J}$, la partie $\widehat{\mathscr{O}}_r$ est fermée en tant qu'intersection de fermés.

Par construction, $\mathscr{O}_r$ est inclus dans $\widehat{\mathscr{O}}_r$, qui est fermé, donc $\overline{\mathscr{O}}_r$ est inclus dans $\widehat{\mathscr{O}}_r$.

Réciproquement, soit A une matrice de rang $k \leqslant r$. D'après le théorème du rang, il existe P et Q inversibles telles que $A = P \mathrm{I}_{m,n,k} Q^{-1}$. Pour $\varepsilon \in \mathbb{K}$,

on définit une matrice par blocs :

$$A(\varepsilon) = P \begin{pmatrix} \mathrm{I}_k & 0 & 0 \\ 0 & \varepsilon \mathrm{I}_{r-k} & 0 \\ 0 & 0 & 0 \end{pmatrix} Q^{-1}.$$

La matrice $A(\varepsilon)$ dépend continûment de ε ; lorsque $\varepsilon \neq 0$, $A(\varepsilon)$ appartient à $\mathscr{O}_r$; pour $\varepsilon = 0$, on a : $A(0) = A$. Ainsi, A appartient à l'adhérence de $\mathscr{O}_r$. Il vient : $\widehat{\mathscr{O}}_r \subset \overline{\mathscr{O}}_r$. □

4.3. Corollaire. *L'unique orbite fermée est l'orbite de la matrice nulle, dite « minimale » :* $\mathscr{O}_0 = \{0\}$.

L'unique orbite ouverte est l'orbite dite « maximale » : $\mathscr{O}_{\min(m,n)}$. *En particulier, si* $m = n$, *le groupe des matrices inversibles est ouvert dans* $\mathscr{M}_n(\mathbb{K})$.

4.4. Remarque. La fonction rg : $\mathscr{M}_{m,n}(\mathbb{K}) \to \mathbb{N}$ n'est pas continue dès que m et n ne sont pas nuls (on munit $\mathbb{N}$ de la topologie discrète ou de la topologie cofinie). En effet, la fibre du fermé $\{r\}$, c'est-à-dire $\mathrm{rg}^{-1}(r) = \mathscr{O}_r$, n'est pas fermée si $r \geqslant 1$. En revanche, la réunion $\mathrm{rg}^{-1}\{0, 1, \ldots, r\} = \bigcup_{0 \leqslant k \leqslant r} \mathscr{O}_k$ est un fermé. On dit que le rang est *semi-continu inférieurement.* Concrètement, on l'illustre à travers le corollaire suivant.

4.5. Corollaire. *Soit* $(A_k)_{k \in \mathbb{N}}$ *une suite de matrices de même rang* r *qui converge vers une matrice* B. *Alors :* $\mathrm{rg}(B) \leqslant r$.

Maintenant que les orbites sont munies d'une topologie, considérons à nouveau la bijection

$$\overline{\alpha_A} : \mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})/G_A \longrightarrow \mathscr{O}_r,$$

où A est une matrice fixée de rang r. Il est raisonnablement naturel de se demander :
- comment mettre une topologie sur le quotient $\mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})/G_A$?
- peut-on espérer que $\overline{\alpha_A}$ soit continue ? voire que ce soit un homéomorphisme, c'est-à-dire que sa bijection réciproque soit aussi continue ?

Ces questions motivent l'introduction de topologie dans les actions, ce qui est l'essentiel du chapitre II.

A. Annexe. Vocabulaire des actions de groupes

Ce paragraphe un peu austère a vocation à rappeler les concepts essentiels pour la manipulation des actions de groupes.

A.1. Définitions

A.1.1. Définition. Soit G un groupe et X un ensemble. On appelle *action à gauche* ou *opération à gauche* ou, plus simplement, *action* ou *opération de G sur X* une application $\alpha : G \times X \to X$, $(g,x) \mapsto g \cdot x$ telle que :

1. $\forall (g,g') \in G,\ \forall x \in X,\ g \cdot (g' \cdot x) = (gg') \cdot x$,
2. $\forall x \in X,\ e \cdot x = x$ (où e est le neutre de G).

A.1.2. Remarque. On associe à l'action α le morphisme $\psi : G \to \mathrm{Bij}(X)$, $g \mapsto \psi_g$ de G vers le groupe $\mathrm{Bij}(X)$ des bijections de X dans X défini par : $\psi_g(x) = g \cdot x$ pour g de G et x de X.

Inversement, étant donné un morphisme $\psi : G \to \mathrm{Bij}(X)$, on définit une action par $\alpha(g,x) = \psi_g(x)$ pour g de G et x de X.

A.1.3. Mise en garde. Ainsi, les données de α et de ψ sont équivalentes, mais il ne faut pas les confondre. En particulier, α n'est pas un morphisme ou, par exemple, une application partielle $G \to X$, $g \mapsto g \cdot x$ non plus (eh ! l'ensemble X n'est pas un groupe en général) ; en tout état de cause, le stabilisateur d'un point x n'est pas un noyau (il n'est pas distingué en général).

Ainsi, si l'on parle du « noyau de l'action », il faut comprendre le noyau de ψ.

A.1.4. Remarque. On définit de même une *action à droite* de G sur X comme étant une application $\beta : G \times X \to X$, $(g,x) \mapsto x \cdot g$, telle que

1. $\forall (g,g') \in G,\ \forall x \in X,\ x \cdot (g \cdot g') = x \cdot (gg')$,
2. $\forall x \in X,\ x \cdot e = x$ (où e est le neutre de G).

On vérifie immédiatement qu'une application $\beta : G \times X \to X$ est une action à droite si et seulement si l'application $\alpha : G \times X \to X$, $(g,x) \mapsto x \cdot g^{-1}$ (c'est-à-dire, $\alpha(g,x) = \beta(g^{-1},x)$ pour tout (g,x)) est une action à gauche.

A.1.5. Définition. Soit $\alpha : G \times X \to X$ une action. *L'orbite* d'un point x de X est l'ensemble noté $\mathscr{O}_x$ ou $\mathrm{Orb}(x)$ ou $G \cdot x$ défini par :

$$G \cdot x = \{g \cdot x\ ,\ \ g \in G\}.$$

L'action est dite *transitive* si elle n'admet qu'une seule orbite : pour tout x de X, $G \cdot x = X$; autrement dit, pour x et x' quelconques dans X, il existe g de G tel que $x' = g \cdot x$.

A.1.6. Définition. Soit G un groupe agissant sur un ensemble X. On appelle ensemble quotient de X par G l'ensemble des orbites de X pour l'action de G. Cet ensemble sera noté X/G.

C'est le moment d'introduire la définition la plus fumeuse de l'ouvrage, mais pourtant bien utile : la notion de *forme normale*. Dans une orbite, on choisira un représentant plus agréable que les autres et on l'appellera forme normale (de l'orbite). Ce choix peut dépendre de plusieurs facteurs : affectivité, simplicité, visibilité seront les critères principaux. Comme on peut le constater, on nage dans le subjectif, il est donc pertinent de parler de normalité. Un petit résumé sur les formes normales pourra être trouvé dans l'annexe II-E.

A.1.7. Définition. Soit $\alpha : G \times X \to X$ une action. Le *stabilisateur* ou *groupe d'isotropie* d'un point x de X est le sous-groupe noté $\mathrm{Stab}_G(x)$ ou encore G_x défini par :

$$\mathrm{Stab}_G(x) = G_x = \{g \in G \ , \ \ g \cdot x = x\}.$$

L'action est dite *libre* si tous les stabilisateurs sont triviaux, *i.e.* réduits au neutre e : pour tout x de X, $G_x = \{e\}$.

A.1.8. Remarque (Importance du stabilisateur). Il y a au moins trois types de raisons pour s'intéresser aux stabilisateurs des points de X : ils permettent d'identifier une orbite à un quotient G/G_x (c'est la raison principale), ils donnent un critère pour qu'une action soit doublement transitive, ils permettent de donner un analogue des « cercles » dans un contexte très général que l'on rencontrera dans le second tome.

A.1.9. Définition (Propriétés d'une action). Soit $\alpha : G \times X \to X$ une action d'un groupe G sur un ensemble X.

L'action est dite :
- *fidèle* si l'intersection des stabilisateurs des points de X est triviale, *i.e.* s'il n'existe pas d'éléments de G autre que e qui fixe tous les points de X, *i.e.* si le noyau du morphisme de l'action est trivial ;
- *libre* si tous les stabilisateurs des points de X sont triviaux ;
- *doublement transitive* si, étant donné deux couples (x_1, x_2) et (y_1, y_2) d'éléments distincts de X, il existe g de G tel que $g \cdot x_1 = y_1$ et $g \cdot x_2 = y_2$;
- *n fois transitive* (où n est un entier naturel non nul) si, étant donné deux n-listes $(x_i)_{1 \leqslant i \leqslant n}$ et $(y_i)_{1 \leqslant i \leqslant n}$ d'éléments deux à deux distincts de X, il existe g de G tel que $g \cdot x_i = y_i$ pour tout i ; on dit qu'elle est *n fois simplement transitive* si, de plus, un tel g est unique ;

– *simplement transitive* si elle est une fois simplement transitive, c'est-à-dire si pour tout x, y de X, il existe un unique g de G tel que $g \cdot x = y$. Notons que dans ce cas, l'application $g \mapsto g \cdot x$ est une bijection de G vers X pour tout x de X.

A.1.10. Exercice. *Montrer qu'une action est doublement transitive si et seulement si la restriction de l'action à chacun des stabilisateurs des points de X est transitive sur $X \setminus \{x\}$*(autrement dit, pour tout x de X, le groupe G_x agit transitivement sur $X \setminus \{x\}$).

A.1.11. Exemple. L'action du groupe linéaire sur les droites est transitive mais pas fidèle : les homothéties agissent trivialement.

Une action induit une action fidèle du quotient du groupe par le noyau de l'action, *i.e.* le noyau du morphisme associé à l'action.

A.1.12. Exemple (*E uno plures*).

Si G agit sur X et n est un entier naturel non nul, on obtient une action dite *diagonale* de G sur X^n par la formule :

$$\forall g \in G,\ \forall (x_1, \ldots, x_n) \in X^n, \quad g \cdot (x_1, \ldots, x_n) = (g \cdot x_1, \ldots, g \cdot x_n).$$

A.1.13. Remarque. Soit $\alpha : G \times X \to X$ une action (à gauche) d'un groupe G sur un ensemble X. Pour tout ensemble Y, on hérite d'une action sur l'ensemble $\mathscr{F}(X, Y)$ des applications de X dans Y comme suit : pour g élément du groupe, f application de X dans Y, on définit $g \cdot f$ par :

$$\forall x \in X, \quad g \cdot f(x) = f(g^{-1} \cdot x).$$

La présence de l'inverse est nécessaire pour avoir une action à gauche et non à droite. [Comparer avec la représentation de permutation du second tome, § X-A.4.]

De plus, si $Y = \mathbb{K}$ est un corps, $\mathscr{F}(X, \mathbb{K})$ est un espace vectoriel (l'espace des fonctions sur X à valeurs dans $\mathbb{K}$), et l'action ainsi définie est linéaire : autrement dit, c'est une *représentation* de G. C'est une idée fondamentale en théorie des représentations : une action géométrique intéressante de G sur X donne lieu à des représentations linéaires intéressantes de G. Elle sera largement illustrée dans notre étude des représentations des sous-groupes finis de $\mathrm{SO}_3(\mathbb{R})$ et $\mathrm{SU}_2(\mathbb{C})$ dans les chapitres X et XI du second volume.

Bien sûr, avec des groupes et des ensembles infinis, il y a trop de fonctions. On ne prend alors que les fonctions continues, ou dérivables, ou polynomiales, ou bien on considère des avatars de fonctions : distributions, sections de fibrés, etc.

A.1.14. Exercice. *On suppose que le groupe G agit sur l'ensemble X et donc sur l'ensemble des fonctions $\mathscr{F}(X,Y)$ comme ci-dessus. On note*

$$\mathscr{F}(X,Y)^G := \{f \in \mathscr{F}(X,Y),\ g \cdot f = f, \forall g \in G\}$$

l'ensemble des fonctions G-invariantes de X dans Y.

1. *Montrer que si f vérifie $g \cdot f = f$ pour un g de G, alors $h \cdot f = f$ pour tout h dans le sous-groupe $\langle g \rangle$ engendré par g.*
2. *En déduire que si f vérifie $g \cdot f = f$, alors elle définit par passage au quotient une fonction $\overline{f}$ de $X/\langle g \rangle$ vers Y telle que $\overline{f}(\overline{x}) = f(x)$, où $\overline{x}$ désigne l'orbite de x. En déduire que l'ensemble des fonctions de X vers Y invariantes par g est en bijection avec $\mathscr{F}(X/\langle g \rangle, Y)$.*

   On considérera l'application qui envoie f sur $\overline{f}$.

   Cet exercice trouvera une application haute en couleurs au moment de l'étude des solides platoniciens, voir l'exercice XII-C.6.

Actions d'un groupe sur lui-même

A.1.15. Définition. Soit G un groupe. On définit trois actions de G sur lui-même :
– par *translations à gauche* : l'action de g sur h est : $g \cdot h = gh$;
– par *translations à droite* : l'action de g sur h est : $g \cdot h = hg^{-1}$;
– par *conjugaison* : l'action de g sur h est : $g \cdot h = ghg^{-1}$.

L'action d'un groupe G par translations est libre et transitive. L'action par conjugaison n'est jamais libre (sauf groupe trivial) et pas nécessairement fidèle : son noyau est le *centre du groupe*, souvent noté $Z(G)$.

A.1.16. Définition. Soit G un groupe et H un sous-groupe de G. On appelle *centralisateur* de H l'intersection des stabilisateurs des éléments de H pour l'action par conjugaison :

$$Z_G(H) = \{g \in G\ ,\ \ \forall h \in H,\ ghg^{-1} = h\}.$$

(Ainsi, par exemple, le centre de G est : $Z(G) = Z_G(G)$.)

Ces actions induisent des actions sur les parties de G : par exemple, si g est un élément de G et V une partie de G, gV est l'ensemble des éléments gh où h parcourt V ; gVg^{-1} est l'ensemble des ghg^{-1} où h parcourt V, etc.

A.1.17. Définition. Soit G un groupe et H un sous-groupe de G. On appelle *normalisateur* de H le stabilisateur de H pour l'action par conjugaison :

$$N_G(H) = \{g \in G\ ,\ \ \forall h \in H,\ ghg^{-1} \in H\}.$$

A.1.18. Exercice. *Montrer que le normalisateur de H dans G est le plus grand sous-groupe de G dans lequel H est distingué.*

On suppose maintenant le groupe fini. En invoquant le théorème 3.2, montrer que le cardinal de l'ensemble des sous-groupes de G qui sont conjugués à H est égal à l'indice de $N_G(H)$ dans G, c'est-à-dire au cardinal de l'ensemble quotient $G/N_G(H)$.

A.1.19. Exercice. *Montrer que le centralisateur d'un sous-groupe est un sous-groupe distingué de son normalisateur. Exemple : Quel est le quotient du normalisateur par H lorsque G est le groupe linéaire et H le groupe des matrices diagonales*[5] *?*

Principe de translation

A.1.20. Principe. *Dans l'action d'un groupe (resp. d'un groupe topologique, d'un groupe de Lie) sur lui-même par translations (à gauche ou à droite), la translation par un élément g est une bijection (resp. un homéomorphisme, un difféomorphisme). Aussi, elle permet d'étendre des propriétés vraies au voisinage du neutre vers un voisinage de g. (Par exemple, le neutre e est un point isolé ou pas de G ; le singleton $\{e\}$ est fermé ; si une partie V est un voisinage [...] du neutre, gV est un voisinage [...] de g, etc.)*

Principe de conjugaison

A.1.21. Principe (D. Perrin [29] – M. Audin [4]). *Si g est un élément d'un groupe de transformations G,*

(i) *un conjugué hgh^{-1} de g par un élément $h \in G$ est « de même nature géométrique » que g ;*

(ii) *les éléments définissant cette « nature » sont, pour le conjugué hgh^{-1}, les images de ceux de g par h.*

A.1.22. Exemples. Voici une batterie d'exemples plus ou moins familiers.

1. Soit $G = \mathfrak{S}_n$ un groupe symétrique. Le conjugué d'un p-cycle est un autre p-cycle, et, plus précisément, si $g = (i_1, \ldots, i_p)$ est un p-cycle, hgh^{-1} est le p-cycle $\big(h(i_1), \ldots, h(i_p)\big)$.
2. Dans une orbite, tous les stabilisateurs sont conjugués. Plus précisément, soit G un groupe opérant sur un ensemble X. Pour $x \in X$, on note G_x son stabilisateur. Alors, $G_{g\cdot x} = gG_xg^{-1}$.
3. Soit ρ une rotation d'axe D et ϕ une isométrie. Alors, $\phi\rho\phi^{-1}$ est encore une rotation, d'axe $\phi(D)$ (et de même angle).

[5] C'est le groupe symétrique qui permute les éléments de la diagonale ! Cet exemple est crucial pour étendre la décomposition de Bruhat (paragraphe I-1 du second volume) aux groupes classiques.

On utilise très souvent le principe de conjugaison dans des situations de commutation : si g commute à h, alors g stabilise les éléments caractéristiques de h (par exemple points fixes, espaces propres...). Le lemme suivant sera utilisé plusieurs fois, sinon dans la lettre, du moins dans l'esprit.

A.1.23. Lemme. *Soit G un sous-groupe du groupe linéaire. Supposons que pour tout vecteur non nul v, il existe un élément h de G dont la droite engendrée par v est un espace propre,* i.e. *: $\mathbb{K}v = \mathrm{Ker}(h - \lambda)$ pour un scalaire λ convenable. Alors, le centre de G est constitué des homothéties qui appartiennent à G. En particulier, le centre de $\mathrm{GL}_n(\mathbb{K})$ est le sous-groupe $\mathbb{K}^*$ des homothéties de $\mathrm{GL}_n(\mathbb{K})$.*

En effet, si g est un élément du centre de G, le principe de conjugaison implique que tout vecteur non nul est propre pour g. Il en résulte que g stabilise toute droite et donc g est une homothétie, par la proposition II-4.3.6.

A.1.24. Exemple. Lorsque G est le groupe linéaire, le lemme s'applique : pour v vecteur non nul, choisissons un hyperplan P ne contenant pas v ; la matrice h qui admet la droite engendrée par v (resp. P) comme espace propre pour valeur propre 1 (resp. 2) est inversible (en caractéristique différente de 2, bien sûr !). On peut adapter l'argument pour le groupe spécial linéaire (sur $\mathbb{R}$ ou $\mathbb{C}$, du moins).

Pour le groupe orthogonal $\mathrm{O}(n)$ ou le groupe unitaire $U(n)$, c'est encore plus facile : pour v vecteur non nul, la réflexion d'hyperplan qui fixe $v^\perp$ convient.

A.2. Étudier une action

Voici un programme générique pour étudier une action d'un groupe G sur un ensemble (ou un espace topologique) :
- classer les orbites (par un ensemble combinatoire assez simple) ;
- trouver une « forme normale » (c'est-à-dire, un élément « plus simple que les autres » dans chaque orbite) ;
- (dans le cadre topologique) décrire les adhérences de chaque orbite.

C'est parfois très facile : pour l'action de $\mathrm{GL}_n(\mathbb{K})$ sur $\mathbb{K}^n$ (où $\mathbb{K}$ est un corps et n un entier naturel), il y a deux orbites : l'ensemble des vecteurs non nuls et le singleton $\{0\}$; comme forme normale, on peut choisir le premier vecteur de la base canonique ; le vecteur nul forme une orbite fermée qui est dans l'adhérence de l'autre orbite.

Lorsque l'action est plus riche, c'est souvent plus compliqué : le théorème d'inertie de Sylvester classe les orbites de formes quadratiques réelles sous l'action du groupe linéaire par leur signature, voir définition V-B.3.5.

B. Développement. Théorème du rang et algorithme du pivot de Gauss

Dans ce chapitre, nous avons vu comment interpréter le théorème du rang en termes d'action de groupe. En *amont* de ce théorème, il y a bien sûr sa preuve. La preuve abstraite est très courte : choisir un supplémentaire du noyau et vérifier qu'il est isomorphe à l'image en y restreignant l'application, puis calculer les dimensions. Dans une autre ligne de pensée, on peut en donner une preuve constructive par l'algorithme de Gauss.

B.1. Proposition. *Soit m et n deux entiers naturels et $\mathbb{K}$ un corps. Par des opérations élémentaires sur les lignes et les colonnes, on peut transformer toute matrice $m\times n$ en la matrice $\mathrm{I}_{m,n,r}$ pour un entier r bien déterminé par la matrice.*

Voici une esquisse de preuve que l'on complétera en exercice. On déduira au passage de la proposition que le rang d'une matrice est égal au rang de sa transposée.

On définit le rang d'une matrice comme la dimension de l'espace engendré par ses colonnes. Il faut montrer qu'il est invariant par des opérations élémentaires sur les rangées. (Si l'on opère sur les colonnes, c'est assez clair.) Il suffit pour cela de constater que le rang est la dimension de l'image de l'application linéaire « multiplication par la matrice », et de savoir que la dimension est préservée par pré- ou post-composition par un isomorphisme (cela résulte aussitôt de la caractérisation : « une bijection envoie une base sur une base », et coûte donc moins cher que le théorème du rang proprement dit). Ce jeu exploite la double interprétation d'une matrice inversible comme matrice de changement de base ou comme matrice d'une bijection.

On peut donc effectuer les opérations prescrites par l'algorithme de Gauss sans crainte de modifier le rang et ramener la matrice initiale à la forme voulue.

B.2. Ubiquité de l'algorithme du pivot de Gauss

Il y a bien d'autres applications de l'algorithme de Gauss et ses variantes, par exemple :

- calcul pratique du rang d'une matrice,
- résolution de systèmes linéaires et réduction à la forme échelonnée (voir le chapitre IV),
- calcul de l'inverse d'une matrice,
- recherche d'un système d'équations d'un sous-espace vectoriel défini par une famille génératrice, recherche d'une base d'un sous-espace vectoriel défini par un système d'équations,

- engendrement de $\mathrm{GL}_n(\mathbb{K})$ par les matrices élémentaires (dilatations, transvections, éventuellement permutations),
- engendrement de $\mathrm{SL}_n(\mathbb{K})$ par les transvections (lequel admet à son tour d'utiles applications : centre, automorphismes, groupe dérivé, simplicité du quotient de $\mathrm{SL}_n(\mathbb{K})$ par ses homothéties) ;
- décomposition de Bruhat (voir le § I-1 du second volume).

En adaptant l'algorithme à un anneau euclidien, ce qui revient essentiellement à subroger le quotient par le pivot a_{i1}/a_{11} à une division euclidienne (« algorithme du pivot de Gauss avec reste », si l'on veut), de nouvelles applications apparaissent :

- recherche effective des invariants de similitude (voir le chapitre III),
- engendrement de $\mathrm{SL}_n(\mathbb{Z})$ par les transvections à coefficients entiers, présentation de $\mathrm{SL}_2(\mathbb{Z})$,
- théorème de la base adaptée et classification des groupes abéliens de type fini (très semblable au précédent).

Osons une formule : *l'algèbre linéaire est une conceptualisation de l'algorithme de Gauss*[6]. Il donne différents résultats si l'on ne s'autorise que certaines opérations élémentaires ou si on l'applique à différents types de matrices (rectangulaires, inversibles...).

Limites de l'algorithme de Gauss : complexité et précision

B.2.1. Proposition. *La complexité de l'algorithme du pivot de Gauss est* $O(n^3)$.

Cela signifie que le nombre d'opérations (multiplications, divisions, additions) sur les coefficients qu'il faut effectuer pour résoudre un système de n équations à n inconnues est de l'ordre de grandeur de n^3 lorsque n est grand.

Sous l'hypothèse que chacune de ces opérations prend un temps constant ou presque, ce qui est le cas si les coefficients sont représentés dans l'ordinateur par des flottants ou des éléments d'un corps fini, cela donne une idée du temps nécessaire à la résolution.

Démonstration. Soit $A = (a_{ij})_{i,j=1,\dots,n}$ la matrice du système. La première étape de l'algorithme du pivot de Gauss consiste à manipuler les lignes pour rendre la matrice triangulaire supérieure. Pour cela, on suppose que $a_{11} \neq 0$ —sinon, $n-1$ comparaisons et une permutation de deux lignes permettent de remplacer le pivot a_{11} par le coefficient de valeur absolue maximale sur la première colonne.

[6] *In vino veritas.* Pour ceux qui n'ont pas fait l'option latin-grec ou qui ne seraient pas œnologues, cela signifie : « Un petit blanc bien tassé donne des visions audacieuses. »

Soit i un entier compris entre 2 à n. On remplace la i^{e} ligne L_i du système par $L_i - \rho_i L_1$, où $\rho_i = a_{i1}/a_{11}$. Pour cela, on effectue une division pour calculer ρ_i, n multiplications et autant d'additions pour la matrice et le second membre, soit $(n+1)$ multiplications et n additions. Comme il y a $n-1$ lignes, cela fait, pour cette étape, n^2-1 multiplications et n^2-n additions. On s'est ramené à un système $(n-1)\times(n-1)$. Pour cette première étape, on trouve donc :
– nombre de multiplications : $\sum_{k=1}^{n}(k^2-1) = n(n-1)(2n+5)/6 \sim n^3/3$,
– nombre d'additions : $\sum_{k=1}^{n} k(k-1) = n(n-1)(n+1)/3 \sim n^3/3$.
On a à présent un système triangulaire à résoudre. La dernière variable s'obtient avec une division. L'avant-dernière s'obtient avec une multiplication, une soustraction et une division. Pour $k \geqslant 1$, le calcul de la variable numéro $n-k$ nécessite $k-1$ multiplications, $k-1$ additions et une division et une soustraction, soit, en tout, $k+1$ multiplications et k additions. Dans cette étape, on trouve donc :
– nombre de multiplications : $\sum_{k=1}^{n}(k+1) = (n+2)(n+1)/2 \sim n^2/2$,
– nombre d'additions : $\sum_{k=1}^{n} k = n(n+1)/2 \sim n^2/2$.
En ajoutant le nombre d'opérations des deux étapes et en supposant n grand, on constate que l'algorithme de Gauss demande environ $n^3/3$ multiplications et autant d'additions ; en simplifiant un peu : *l'algorithme de Gauss pour une matrice $n\times n$ demande environ $2n^3/3$ opérations.* □

B.2.2. Remarque. On a bien fait de négliger les comparaisons, car il en faut au maximum $n-1$ pour la première colonne, $n-2$ pour la deuxième, etc., soit $O(n^2)$ en tout. Quant aux permutations de lignes, il y en a $O(n)$ au maximum.

B.2.3. Remarque (Un système de grande taille). On veut résoudre numériquement une équation aux dérivées partielles pour décrire, disons,un fluide dans un cube de côté 1 mètre. On veut appliquer une méthode comme le schéma d'Euler implicite. On linéarise l'équation et l'on discrétise l'espace et le temps : disons que l'on coupe chaque arête du cube en morceaux de 1 cm. La solution de l'EDP est décrite par un réel pour chaque petit cube.

Alors, pour chaque pas de temps, on est amené à résoudre un système avec : $n = 100^3 = 10^6$ inconnues et autant d'équations. Avec l'algorithme de Gauss, il faudra donc effectuer $2(10^6)^3/3 = 0,666\times10^{18}$ opérations élémentaires pour chaque pas de temps. La fréquence du processeur des micro-ordinateurs d'aujourd'hui est de l'ordre du gigahertz (GHz) ; si l'on compte sur 1 GFLOPS, soit[7] 10^9 opérations sur les flottants par seconde, la résolution d'un seul système prendra... plus de vingt ans !

[7]L'acronyme FLOPS signifie *FLoating Operations Per Second.*

C'est évidemment inconcevable, d'où la nécessité de trouver des méthodes (beaucoup) plus efficaces pour résoudre les systèmes linéaires. C'est un thème de recherche actuel.

B.2.4. Exemple (Problèmes de conditionnement). Aux problèmes de temps de calcul s'ajoutent des problèmes de précision des calculs : si une matrice est *mal conditionnée*, l'algorithme de Gauss risque de faire apparaître des pivots très petits par lesquels il va falloir diviser. Si l'on calcule avec des flottants, cela conduit à des pertes de précision irrémédiables ; si l'on fait des calculs exacts avec des rationnels, on se retrouve à manipuler des dénominateurs gigantesques, d'où une explosion du temps de chaque opération élémentaire sur les coefficients et des problèmes de mémoire. Là encore, il s'agit d'un thème de recherche très actuel.

Référence. Pour une présentation plus élaborée mais très accessible de ces problématiques numériques, voir par exemple le livre [32] de Michelle Schatzman.

C. Exercices

C.1. Exercice (Rang d'une matrice entière). Soit A une matrice à coefficients dans $\mathbb{Z}$. Comparer le rang de A sur $\mathbb{Q}$, et sur $\mathbb{C}$. Faire aussi une comparaison avec le rang de la réduction $\overline{A}$ de A modulo p.

Automorphismes des algèbres de matrices

C.2. Exercice (Théorème de Skolem-Noether). Soit n un entier naturel et $\mathbb{K}$ un corps. Partant de l'idée qu'il existe (à isomorphisme près) un seul espace vectoriel par dimension, on prouve que tout automorphisme de l'algèbre de matrices $\mathscr{M}_n(\mathbb{K})$ est intérieur.
Soit φ un automorphisme d'algèbres de $\mathscr{M}_n(\mathbb{K})$. Pour i et j entiers entre 1 et n, on note E_{ij} la matrice dont le seul coefficient non nul est un 1 placé en position (i,j) et l'on pose $E'_{ij} = \varphi(E_{ij})$.

1. Prouver que les matrices E'_{ii} sont des projecteurs de rang 1 pour tout i et que leurs images sont en somme directe.

   On pourra utiliser le fait que la somme des E_{ii} est l'identité.

2. Soit e'_1 une base de l'image de E'_{11} et $e'_i = E'_{i1}e'_1$ pour $2 \leqslant i \leqslant n$. Prouver que $\mathbf{e}' = (e'_1, \ldots, e'_n)$ est une base de $\mathbb{K}^n$.

   La relation $E_{ii} = E_{i1}E_{11}$ peut aider à montrer que la famille est génératrice.

3. Soit P la matrice de passage de la base canonique à la base $\mathbf{e}'$. Montrer que φ est la conjugaison par P : pour toute matrice A, $\varphi(A) = PAP^{-1}$. *Calculer l'action de E'_{ij} et de $PE_{ij}P^{-1}$ sur la base $\mathbf{e}'$.*

C.3. Exercice (Idéaux et représentations des algèbres de matrices). Soit n un entier et $\mathbb{K}$ un corps. On va montrer que la théorie des représentations de l'algèbre des matrices $\mathscr{A} = \mathscr{M}_n(\mathbb{K})$ est particulièrement simple. On note E_{ij} la matrice, dite *matrice élémentaire* dont tous les coefficients sont nuls, sauf celui d'indice (i,j) qui vaut 1.

1. Soit $\mathscr{I}$ un idéal bilatère[8] de $\mathscr{A}$. Montrer que $\mathscr{I}$ est égal à $\{0\}$ ou $\mathscr{A}$. *Si A est une matrice non nulle dans I, soit (i,j) l'indice d'un coefficient non nul de A. Vérifier que la matrice élémentaire E_{ij} appartient à $\mathscr{I}$, puis que toutes les matrices élémentaires appartiennent à $\mathscr{I}$, puis que $\mathscr{I} = \mathscr{A}$.*

2. Soit $\mathscr{I}$ un idéal à gauche de $\mathscr{A}$. Soit I l'ensemble des entiers $k \in \{1,\dots,n\}$ pour lesquels il existe un élément A de $\mathscr{I}$ dont la k^{e} colonne n'est pas nulle. Prouver que l'on a : $\mathscr{I} = \bigoplus_{k\in I} \mathscr{A}E_{kk}$. Déduire de cette écriture une description des $\mathscr{A}$-modules $\mathscr{I}$ et $\mathscr{A}/\mathscr{I}$.

   Soit $A = \sum_{i,j} a_{ij}E_{ij}$ un élément de $\mathscr{I}$ dont la k^{e} colonne n'est pas nulle. Vérifier que $\mathscr{I}$ contient une matrice E_{ik} pour i convenable, puis que $\mathscr{I}$ contient $\mathscr{A}E_{kk}$. Inversement, vérifier que $\mathscr{A}E_{kk}$ est l'espace des matrices dont toutes les colonnes sont nulles sauf peut-être la k^{e}.

3. Soit M un $\mathscr{A}$-module qui est un espace vectoriel de dimension finie sur $\mathbb{K}$. Prouver que M est isomorphe à une somme directe de sous-modules isomorphes à $\mathbb{K}^n$. En particulier, tout $\mathscr{A}$-module est semi-simple et $\mathscr{A}$ possède un seul module simple à isomorphisme près.

   Une structure de $\mathscr{A}$-module est définie par un morphisme d'algèbres φ de $\mathscr{A}$ dans l'algèbre des endomorphismes de M. Soit, pour i et j entiers compris entre 1 et n, $e_{ij} = \varphi(E_{ij})$. Vérifier que les e_{jj} sont des projecteurs et que leurs images sont des supplémentaires. Montrer que e_{ij} se restreint en une injection de $e_{jj}M$ dans $e_{ii}M$ et que le rang r de e_{jj} est indépendant de j. On fixe $\left(v_1^{(k)}\right)_{1\leqslant k\leqslant r}$ une base de $e_{11}M$. Montrer que $\left(e_{j1}v_1^{(k)}\right)_{1\leqslant j\leqslant n, 1\leqslant k\leqslant r}$ est une base de M et conclure.

[8] C'est-à-dire un sous-groupe stable par multiplication à gauche et à droite par une matrice quelconque. De même, un idéal à gauche est un sous-groupe stable par multiplication à gauche.

Automorphismes des groupes linéaires

C.4. Exercice (Un automorphisme extérieur). Contrairement aux algèbres de matrices (voir l'exercice C.2), les groupes linéaires possèdent (presque tous) des automorphismes extérieurs, c'est-à-dire qui ne peuvent pas s'écrire sous la forme $g \mapsto PgP^{-1}$ pour P fixé.
Soit n un entier naturel. Considérons l'application

$$\omega : \mathrm{GL}_n(\mathbb{K}) \longrightarrow \mathrm{GL}_n(\mathbb{K}), \quad P \mapsto {}^tP^{-1}.$$

1. Montrer que ω est un automorphisme de $\mathrm{GL}_n(\mathbb{K})$.
2. Montrer que la matrice de passage entre les bases duales est la transposée de l'inverse de la matrice de passage.

   Formulation plus précise. Soit E un espace vectoriel de dimension finie n non nulle, soit $\mathbf{e} = (e_1, \ldots, e_n)$ et $\mathbf{e}' = (e'_1, \ldots, e'_n)$ deux bases de E. On note $P = P_{\mathbf{e},\mathbf{e}'}$ la matrice de passage : pour $j \in \{1, \ldots, n\}$, la j^{e} de P est la colonne des coordonnées de e'_j dans $\mathbf{e}$. Enfin, on note $\mathbf{e}^* = (\ell_1, \ldots, \ell_n)$ et $\mathbf{e}'^* = (\ell'_1, \ldots, \ell'_n)$ les bases duales respectives de $\mathbf{e}$ et $\mathbf{e}'$: on a par exemple : $\ell_i(e_j) = \delta_{ij}$ (le delta de Kronecker) pour tout couple (i, j).

   Montrer que la matrice de passage de $\mathbf{e}^*$ à $\mathbf{e}'^*$ est ${}^tP^{-1}$.

3. Sauf mention du contraire, on suppose que la caractéristique de $\mathbb{K}$ n'est pas 2. On veut montrer que ω n'est pas un automorphisme intérieur.
   On note $\mathrm{O}(n, \mathbb{K})$ l'ensemble des matrices $g \in \mathrm{GL}_n(\mathbb{K})$ telles que ${}^tgg = \mathrm{I}_n$.
   On suppose qu'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ tel que

   $$\forall g \in \mathrm{GL}_n(\mathbb{K}), \quad {}^tg^{-1} = PgP^{-1}.$$

   (a) Montrer que $P \in \mathrm{O}(n, \mathbb{K})$.

   (b) Montrer que P appartient au centre de $\mathrm{O}(n, \mathbb{K})$: pour tout élément g de $\mathrm{O}(n, \mathbb{K})$, on a : $gP = Pg$.

   (c) En utilisant les matrices diagonales dont tous les coefficients valent 1 sauf un, qui vaut -1, montrer que P est diagonale. En utilisant les matrices diagonales par blocs dont les blocs diagonaux sont

   $$\mathrm{I}_k, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \mathrm{I}_{n-k-2}$$

   pour k entre 0 et $n-2$, montrer que P est un multiple de l'identité.
   On montrera que les espaces propres associés à la valeur propre -1 de toutes ces matrices sont également des espaces propres pour P.

   (d) Montrer que lorsque $\mathbb{K} = \mathbb{F}_2$, l'automorphisme ω de $\mathrm{GL}_2(\mathbb{K})$ est intérieur.

(e) (Plus difficile.) Étudier les automorphismes de $\mathrm{GL}_n(\mathbb{K})$ lorsque $\mathbb{K}$ est de caractéristique 2.

4. Montrer que les groupes suivants que vous connaissez sont stables par la transposition, et donc par $\omega : \mathrm{U}_n(\mathbb{C})$, $\mathrm{O}(n, \mathbb{K})$, $\mathrm{O}_{p,q}(\mathbb{R})$, où $\mathbb{K}$ est un corps quelconque et n, p, q sont des entiers naturels non nuls.

Chapitre II

Groupes topologiques, actions continues

Loki : I have an army.
Tony Stark : We have a Hulk.

The Avengers, Joss Whedon, 2012.

Dans ce chapitre, on introduit de la topologie dans des situations d'actions de groupes. On en verra des applications dans tout le livre, que ce soit en algèbre linéaire ou en géométrie. On a en vue des démonstrations courtes de certains théorèmes (voir, pour commencer, le paragraphe 2 en algèbre linéaire) mais, surtout, l'étude de certains objets algébriques en tant qu'objets géométriques (par exemple, l'ensemble des formes quadratiques à homothétie près sur $\mathbb{R}^2$). La topologie joue alors le rôle d'une structure intermédiaire, un peu plus rigide que les ensembles sans structure mais plus facile à manier que celle de variétés (différentielles, que l'on verra au chapitre IX, ou algébriques, sujet que l'on effleurera prudemment, par exemple au § I-2 du second tome).

Avant d'attaquer de front les problèmes d'actions topologiques, nous engageons le lecteur à lire l'annexe I-A pour les définitions et propriétés de base sur les actions de groupes. Dans tout le chapitre, $\mathbb{K}$ désignera le corps des réels ou des complexes. La topologie des espaces de matrices sur $\mathbb{K}$ est donnée en annexe A du présent chapitre.

1. Groupes topologiques : définitions et exemple fondamental

Pour toutes les définitions et autres résultats de base sur la topologie, nous invitons le lecteur à se rapporter aux différents appendices.

1.1. Définition. Un *groupe topologique* est un groupe G muni d'une topologie pour laquelle sont continues les applications de produit et de passage à l'inverse :

$$\begin{array}{rcl} \mu : G \times G & \longrightarrow & G \\ (g,h) & \longmapsto & gh \end{array} \quad \text{et} \quad \begin{array}{rcl} \iota : G & \longrightarrow & G \\ g & \longmapsto & g^{-1}. \end{array}$$

1.2. Remarque. Comme l'inversion ι est involutive, *i.e.* $\iota^2 = \mathrm{Id}_G$, c'est un homéomorphisme.

Un outil simple et fondamental dans l'étude des groupes topologiques est le *principe de translation*, selon lequel une propriété vraie dans un voisinage du neutre a des chances de rester vraie dans un voisinage de n'importe quel point. Il repose sur le lemme suivant, qui sera implicitement utilisé en permanence.

1.3. Lemme. *Soit G un groupe topologique et soit g un élément de G. Les translations à gauche et à droite*

$$L_g : G \to G,\ h \mapsto gh \quad \text{et} \quad R_g : G \to G,\ h \mapsto hg$$

sont des homéomorphismes.

Démonstration. L'application L_g (resp. R_g) est continue et bijective et son inverse est $L_{g^{-1}}$ (resp. $R_{g^{-1}}$), qui est continue. □

On comprendra sans peine que le groupe linéaire joue, pour les espaces vectoriels, le même rôle que le groupe des permutations pour les ensembles : ce sont les groupes formés par morphismes inversibles qui préservent la structure. L'exemple fondamental (de ce livre) est donc le suivant.

1.4. Proposition. *Soit n un entier naturel. Si $\mathbb{K}$ est $\mathbb{R}$ ou $\mathbb{C}$, le groupe $\mathrm{GL}_n(\mathbb{K})$, muni de la topologie induite par la topologie (normique) de $\mathscr{M}_n(\mathbb{K})$, est un groupe topologique. De plus, $\mathrm{GL}_n(\mathbb{K})$ est ouvert et dense dans $\mathscr{M}_n(\mathbb{K})$. Le groupe $\mathrm{GL}_n(\mathbb{C})$ est connexe par arcs, mais $\mathrm{GL}_n(\mathbb{R})$ n'est pas connexe.*

Démonstration. Le produit des matrices est continu, car les composantes du produit sont des polynômes en les coefficients des matrices. L'inversion est continue sur $\mathrm{GL}_n(\mathbb{K})$. En effet, pour toute matrice carrée $A = (a_{ij})$, notons $\mathrm{com}(A)$ la *comatrice* de A, c'est-à-dire la matrice des cofacteurs (m_{ij}) avec $m_{ij} = (-1)^{i+j}\Delta_{ij}$, Δ_{ij} étant le mineur associé à la ligne i et la colonne j de A. Par la formule de Cramer, on a pour toute matrice A

de $\mathrm{GL}_n(\mathbb{K})$: $A^{-1} = {}^t\mathrm{com}(A)/\det(A)$: le passage à l'inverse est donc le quotient de deux fonctions polynomiales.

Le groupe $\mathrm{GL}_n(\mathbb{K})$ est ouvert dans $\mathscr{M}_n(\mathbb{K})$, car c'est $\det^{-1}(\mathbb{K}^*)$, image réciproque d'un ouvert de $\mathbb{K}$ par l'application continue déterminant.

En reprenant les notations du chapitre I, on observe que $\mathrm{GL}_n(\mathbb{K})$ est l'orbite $\mathscr{O}_n$ des matrices de rang n. Son adhérence est, par la proposition I-4.1 :

$$\overline{\mathrm{GL}_n(\mathbb{K})} = \overline{\mathscr{O}_n} = \bigcup_{r=0}^{n} \mathscr{O}_r = \mathscr{M}_n(\mathbb{K}).$$

Supposons que $\mathbb{K} = \mathbb{R}$. Comme le déterminant est une surjection sur $\mathbb{R}^*$ et que $\mathbb{R}^*$ n'est pas connexe, $\mathrm{GL}_n(\mathbb{R})$ n'est pas connexe. [On verra que ses composantes connexes sont les images réciproques $\det^{-1}(\mathbb{R}^{+*})$ et $\det^{-1}(\mathbb{R}^{-*})$.]

Supposons que $\mathbb{K} = \mathbb{C}$: montrons que $\mathrm{GL}_n(\mathbb{C})$ est connexe par arcs. Soit A et B de $\mathrm{GL}_n(\mathbb{C})$. La fonction polynôme $P : \mathbb{C} \to \mathbb{C}$, $z \mapsto \det\big(zA+(1-z)B\big)$ n'est pas nulle puisque l'on a : $P(1) = \det A \neq 0$. Par suite, P a un nombre fini de racines et la partie $\mathscr{C} = \mathbb{C} \setminus P^{-1}(0)$ est connexe par arcs. Soit $\gamma : [0,1] \to \mathscr{C}$ une application continue telle que $\gamma(0) = 0$ et $\gamma(1) = 1$. Alors, l'arc $[0,1] \to \mathscr{M}_n(\mathbb{C})$, $t \mapsto \gamma(t)A + \big(1-\gamma(t)\big)B$ est à valeurs dans $\mathrm{GL}_n(\mathbb{C})$ et relie A et B. La connexité par arcs de $\mathrm{GL}_n(\mathbb{C})$ en résulte. □

Les « groupes classiques » que nous étudierons tout au long du livre sont des sous-groupes de $\mathrm{GL}_n(\mathbb{K})$, pour n convenable, d'où son statut d'exemple fondamental. On peut y voir une vague analogie avec le fait que tout groupe fini est isomorphe à un sous-groupe d'un groupe symétrique.

Voici les premiers exemples de groupes classiques (liste non exhaustive des groupes apparaissant dans le livre). Comme d'habitude, ${}^tA$ est la transposée de la matrice A et A^* est la transposée de sa conjuguée (coefficient par coefficient).

1.5. Définition (Exemples de groupes classiques). Soit n un entier naturel et $\mathbb{K}$ un corps. On définit les groupes suivants (chacun étant « le groupe [...] d'ordre n » ; si $n = 0$, le groupe est réduit à $\{1\}$) :

groupe spécial linéaire : $\mathrm{SL}_n(\mathbb{K}) = \{A \in \mathrm{GL}_n(\mathbb{K}) \ , \ \det A = 1\}$,

groupe orthogonal : $\mathrm{O}(n,\mathbb{K}) = \{A \in \mathrm{GL}_n(\mathbb{K}) \ , \ {}^tA = A^{-1}\}$,

groupe unitaire : $\mathrm{U}(n,\mathbb{C}) = \{A \in \mathrm{GL}_n(\mathbb{K}) \ , \ A^* = A^{-1}\}$,

groupe spécial orthogonal : $\mathrm{SO}(n,\mathbb{K}) = \mathrm{O}(n,\mathbb{K}) \cap \mathrm{SL}_n(\mathbb{K})$,

groupe spécial unitaire : $\mathrm{SU}(n,\mathbb{C}) = \mathrm{U}_n(\mathbb{C}) \cap \mathrm{SL}_n(\mathbb{C})$.

On pourra aussi trouver les notations $\mathrm{O}_n(\mathbb{K})$, $\mathrm{SO}_n(\mathbb{K})$, ou si le contexte s'y prête O_n, SO_n, U_n, SU_n !

2. Premières applications de la topologie en algèbre linéaire

2.1. Polynôme caractéristique d'un produit

2.1.1. Proposition. *Soit A et B deux matrices carrées de même taille à coefficients réels ou complexes. Alors, AB et BA ont le même polynôme caractéristique.*

Démonstration. Notons n la taille des matrices et $\mathbb{K}$ le corps ($\mathbb{R}$ ou $\mathbb{C}$).

Si A est inversible, on a : $AB = A(BA)A^{-1}$, si bien que AB et BA sont semblables. Pour tout scalaire λ, on a alors :

$$\det(AB - \lambda \operatorname{Id}) = \det A \det(BA - \lambda \operatorname{Id}) \det A^{-1}.$$

Comme $\mathbb{K}$ est infini, A et B ont donc le même polynôme caractéristique.

À présent, soit A quelconque. Grâce à la densité du groupe linéaire dans l'espace des matrices, que l'on a vue dans la proposition 1.4, il existe une suite $(A_k)_{k\in\mathbb{N}}$ de matrices inversibles qui converge vers A. Or, l'application « polynôme caractéristique » $\chi : \mathscr{M}_n(\mathbb{K}) \to \mathbb{K}_n[\lambda]$, $M \mapsto \chi_M$, est continue de l'espace des matrices vers l'espace des polynômes de degré inférieur à n, car chacun des coefficients de $\chi(M)$ est une fonction polynomiale des coefficients de χ_M (par exemple, le coefficient de degré $n-1$ est l'opposé de la trace, celui de degré 0 est le déterminant ; voir la proposition II-A.1.2 et l'exercice « Mineurs et polynôme caractéristique » dans le paragraphe I-2.2 du second tome pour d'autres formules). Grâce à la continuité et au début de la preuve, il vient :

$$\chi_{AB} = \lim_{k\to+\infty} \chi_{A_kB} = \lim_{k\to+\infty} \chi_{BA_k} = \chi_{BA}.$$

□

2.2. Densité des matrices diagonalisables et théorème de Cayley-Hamilton

On rappelle qu'une matrice A de $\mathscr{M}_n(\mathbb{K})$ est diagonalisable si elle est semblable à une matrice diagonale. Nous pénétrons ici dans le vaste programme de la réduction qui sera largement étudié dans le chapitre III. Voici donc un avant-goût.

Pour n un entier naturel, on note $\mathscr{M}_n^{\mathrm{reg}}(\mathbb{C})$ l'ensemble des matrices $n \times n$ ayant n valeurs propres distinctes (matrices *régulières*) et $\mathscr{D}_n(\mathbb{C})$ l'ensemble des matrices $n \times n$ diagonalisables. Rappelons que comme une matrice régulière a un polynôme minimal scindé à racines simples, elle est diagonalisable : $\mathscr{M}_n^{\mathrm{reg}}(\mathbb{C}) \subset \mathscr{D}_n(\mathbb{C})$.

2.2.1. Proposition. *Soit n un entier naturel. Alors, $\mathscr{M}_n^{\mathrm{reg}}(\mathbb{C})$ et,* a fortiori, *$\mathscr{D}_n(\mathbb{C})$, sont denses dans $\mathscr{M}_n(\mathbb{C})$.*

Démonstration. Soit A une matrice $n \times n$. Comme $\mathbb{C}$ est algébriquement clos, le polynôme caractéristique de A a n racines et on peut trigonaliser A : il existe P inversible telle que $T = P^{-1}AP$ soit triangulaire supérieure. On va construire une suite de matrices de $\mathscr{M}_n^{\mathrm{reg}}(\mathbb{C})$ qui converge vers T.

Soit $\lambda_1, \ldots, \lambda_n$ les coefficients diagonaux de T. Soit k un entier naturel. On pose $\varepsilon_1^k = 0$ et l'on choisit de proche en proche des complexes $\varepsilon_2^k, \ldots, \varepsilon_n^k$ de sorte que, pour tout $i \geqslant 2$, on ait :

$$|\varepsilon_i^k| < 2^{-k} \quad \text{et} \quad \lambda_i + \varepsilon_i^k \notin \{\lambda_1 + \varepsilon_1^k, \ldots, \lambda_{i-1} + \varepsilon_{i-1}^k\}.$$

Notons enfin D_k la matrice diagonale dont les coefficients diagonaux sont $(\varepsilon_1^k, \ldots, \varepsilon_n^k)$. Alors, la suite $(T + D_k)_{k\in\mathbb{N}}$ converge vers T et ses termes appartiennent à $\mathscr{M}_n^{\mathrm{reg}}(\mathbb{C})$.

Par continuité de l'application $M \mapsto PMP^{-1}$, la suite $\left(P(T+D_k)P^{-1}\right)_{k\in\mathbb{N}}$ converge vers $PTP^{-1} = A$ et ses termes, comme les matrices $T + D_k$, appartiennent à $\mathscr{M}_n^{\mathrm{reg}}(\mathbb{C})$. Ainsi, $\mathscr{M}_n^{\mathrm{reg}}(\mathbb{C})$ est dense dans $\mathscr{M}_n(\mathbb{C})$. Comme toute matrice à spectre simple est diagonalisable, la densité de $\mathscr{D}_n(\mathbb{C})$ en résulte. □

2.2.2. Remarque. Le *discriminant* du polynôme caractéristique, voir exercice A.1.4, fournit une fonction polynomiale non nulle $\mathrm{disc} : \mathscr{M}_n(\mathbb{C}) \to \mathbb{C}$ et l'on a : $\mathscr{M}_n^{\mathrm{reg}}(\mathbb{C}) = \mathrm{disc}^{-1}(\mathbb{C}^*)$, si bien que $\mathscr{M}_n^{\mathrm{reg}}(\mathbb{C})$ est un ouvert. On peut en déduire une autre preuve de la densité.

2.2.3. Remarque. Attention ! Il y a un obstacle au passage de $\mathbb{C}$ à $\mathbb{R}$. L'ensemble des matrices 2×2 réelles diagonalisables n'est pas dense dans $\mathscr{M}_2(\mathbb{R})$. Par exemple, au voisinage de la matrice d'une rotation d'angle ni nul, ni plat, le polynôme caractéristique n'est pas scindé. En fait, l'adhérence de l'ensemble des matrices diagonalisables est l'ensemble des matrices trigonalisables (*i.e.* dont le polynôme caractéristique est scindé sur les réels). Voir pour cela l'exercice III-D.7.

2.2.4. Proposition (Cayley-Hamilton). *Soit A une matrice carrée réelle ou complexe et χ_A son polynôme caractéristique. Alors : $\chi_A(A) = 0$.*

Résumé. L'égalité $\chi_A(A) = 0$ est triviale si A est diagonale, donc, par conjugaison, elle reste vraie si A est diagonalisable, et enfin elle reste vraie pour tout A par densité. □

Démonstration. Soit n la taille de la matrice. L'application « polynôme caractéristique » $\chi : \mathscr{M}_n(\mathbb{C}) \to \mathbb{C}_n[\lambda]$ est continue par la proposition A.1.2. Soit A un élément de $\mathscr{M}_n^{\mathrm{reg}}(\mathbb{C})$ et soit $\lambda_1, \ldots, \lambda_n$ ses valeurs propres. D'une part, on a : $\chi_A(\lambda) = \prod_{i=1}^n (\lambda - \lambda_i)$. D'autre part, A est diagonalisable : il existe une matrice inversible P telle que $D = P^{-1}AP$ soit diagonale. Les

polynômes caractéristiques de A et D sont égaux, et il est immédiat de voir que l'on a : $\chi_A(D) = 0$. Par suite, on a : $\chi_A(A) = P\chi_A(D)P^{-1} = 0$.
Enfin, les coefficients de la matrice $\chi_A(A)$ sont des polynômes en les coefficients de A : en effet, si l'on note $\chi_A(\lambda) = \sum_{i=0}^{n} c_i(A)\lambda^i$, les $c_i(A)$ sont des polynômes en les coefficients de A et $\chi_A(A) = \sum_{i=0}^{n} c_i(A)A^i$. En particulier, l'application $f : \mathscr{M}_n(\mathbb{C}) \to \mathscr{M}_n(\mathbb{C})$, $A \mapsto \chi_A(A)$ est continue[1]. Comme sa restriction à $\mathscr{D}_n(\mathbb{C})$ est nulle et que $\mathscr{D}_n(\mathbb{C})$ est dense dans $\mathscr{M}_n(\mathbb{C})$, f est partout nulle. □

2.3. Connexité de quelques parties liées au rang

2.3.1. Proposition. *Soit n un entier naturel et $r \in \{0, \dots, n\}$.*

(i) *L'ensemble $\mathscr{O}_r$ des matrices complexes $n \times n$ de rang r est connexe.*

(ii) *L'ensemble $\mathscr{P}_r$ des projecteurs de $\mathscr{M}_n(\mathbb{C})$ de rang r est connexe.*

(iii) *Les composantes connexes de l'ensemble des projecteurs de $\mathscr{M}_n(\mathbb{C})$ sont les $\mathscr{P}_r$ $(0 \leqslant r \leqslant n)$.*

Démonstration. (i) On a déjà noté $\mathrm{I}_{n,r}$ la matrice $n \times n$ diagonale par blocs dont les blocs diagonaux sont I_r, l'identité dans $\mathscr{M}_r(\mathbb{C})$, et la matrice nulle de taille $(n-r)\times(n-r)$. On a alors vu que $\mathscr{O}_r$ est l'image de l'application $\mathrm{GL}_n(\mathbb{C}) \times \mathrm{GL}_n(\mathbb{C}) \to \mathscr{M}_n(\mathbb{C})$, $(P, Q) \mapsto P\mathrm{I}_{n,r}Q^{-1}$. Comme cette application est continue (coefficients polynomiaux en les coefficients de P et Q) et que $\mathrm{GL}_n(\mathbb{C})$ est connexe (proposition 1.4), $\mathscr{O}_r$ est connexe comme image d'un connexe par une application continue.
(ii) Montrons[2] que $\mathscr{P}_r$ est l'orbite de $\mathrm{I}_{n,r}$, c'est-à-dire l'image de l'application $\psi_r : \mathrm{GL}_n(\mathbb{C}) \to \mathscr{M}_n(\mathbb{C})$, $P \mapsto P\mathrm{I}_{n,r}P^{-1}$.
Un projecteur est une matrice A telle que $A^2 = A$. Pour P inversible, on a : $(P\mathrm{I}_{n,r}P^{-1})^2 = P\mathrm{I}_{n,r}P^{-1}$ et, bien sûr : $\operatorname{rg} P\mathrm{I}_{n,r}P^{-1} = r$, si bien que l'image de ψ est contenue dans $\mathscr{P}_r$. Inversement, soit A un projecteur de rang r. Comme A annule le polynôme scindé $X^2 - X$, A est diagonalisable et a pour valeurs propres 1 et 0 : le rang r de A est la multiplicité de la valeur propre 1, si bien que A est semblable à $\mathrm{I}_{n,r}$.
Alors, comme ψ_r est continue et que $\mathrm{GL}_n(\mathbb{C})$ est connexe, son image $\mathscr{P}_r$ est connexe.
(iii) La trace est une application continue qui, restreinte aux projecteurs, coïncide avec le rang[3]. Les parties $\mathscr{P}_r$ $(0 \leqslant r \leqslant n)$ sont connexes et de

[1] Pour décomposer un peu, on peut dire que c'est la composée des applications continues $\mathscr{M}_n(\mathbb{K}) \to \mathbb{K}_n[\lambda] \times \mathscr{M}_n(\mathbb{K})$, $A \mapsto (\chi_A, A)$ et $\mathbb{K}_n[\lambda] \times \mathscr{M}_n(\mathbb{K}) \to \mathscr{M}_n(\mathbb{K})$, $(P, M) \mapsto P(M)$.

[2] Cela consiste à dire en termes savants le fait bien connu qu'un projecteur a, dans une base convenable, une matrice diagonale dont les coefficients diagonaux valent 0 ou 1.

[3] On peut y voir une émanation du principe de conjugaison : trace et rang sont invariants par conjugaison et prennent la même valeur en les projecteurs $\mathrm{I}_{n,r}$.

plus fermées pour la topologie restreinte aux projecteurs. Comme il y en a un nombre fini, les composantes connexes de l'ensemble des projecteurs sont des réunions de ces parties, voir annexe B.10. Mais, si r et s sont deux entiers distincts, $\mathscr{P}_r \cup \mathscr{P}_s$ n'est pas connexe puisque la trace fournit une application continue à valeurs dans $\{r, s\}$, non constante sur $\mathscr{P}_r \cup \mathscr{P}_s$. □

3. Action continue, théorème d'homéomorphisme

Dans le chapitre I, on a vu la pertinence des actions de groupes pour les problèmes d'algèbre linéaire ; le paragraphe 2 montre la valeur ajoutée de la topologie pour ces questions ou, plus précisément, celle des actions continues de groupes topologiques.

Pour mélanger les deux aspects, nous allons décrire le formalisme des actions continues de groupes topologiques.

En particulier, étant donné une action $\alpha : G \times X \to X$ et un point x de X, le théorème I-3.2 établit une bijection $\overline{\alpha_x} : G/G_x \to G \cdot x$. Nous voulons à présent mettre un peu de topologie dans ces considérations. La question suivante est très naturelle : si G et X sont des espaces topologiques, quelles hypothèses raisonnables imposer à l'action pour que $\overline{\alpha_x}$ soit un homéomorphisme ?

3.1. Actions continues

3.1.1. Définition. Soit G un groupe topologique et X un espace topologique. On appelle *action continue* (ou *opération continue*) de G sur X une action $\alpha : G \times X \to X$ qui est une application continue[4].

3.1.2. Convention. On supposera implicitement que G et X sont des espaces séparés, voir l'annexe A.3.

3.1.3. Remarque. Soit $\alpha : G \times X \to X$ une action continue. Alors :
- pour tout x de X, l'application α_x est continue, car c'est la composée $\alpha_x : G \to G \times \{x\} \xrightarrow{\alpha} X$;
- pour tout x de X, le stabilisateur G_x de x est fermé car :

$$G_x = \alpha_x^{-1}(\{x\});$$

- pour tout g de G l'application $\tau_g : X \to X$, $x \mapsto \alpha(g, x)$ est un homéomorphisme, car elle est continue comme composée de $x \mapsto (g, x)$ et de α, et sa réciproque est $\tau_{g^{-1}}$.

Pour donner une première interaction entre action et topologie et pour référence ultérieure, voyons une application du « principe de translation ».

[4]Ici, comme implicitement partout dans la suite, la topologie sur $G \times X$ est la topologie produit.

3.1.4. Lemme. *Soit G un groupe topologique qui agit continûment sur un espace topologique X. L'adhérence d'une G-orbite est stable par G, c'est donc une réunion de G-orbites.*

Démonstration. Soit $\mathcal{O}$ une G-orbite et x un point adhérent à $\mathcal{O}$. Soit $g \in G$. Soit W un ouvert contenant $g \cdot x = \tau_g(x)$ (notation τ_g de la remarque ci-dessus). Alors, $V = g^{-1} \cdot W = \tau_{g^{-1}}(W)$ est un ouvert contenant x, donc V rencontre $\mathcal{O}$ en un point y. Mais alors, $g \cdot y$ appartient à l'orbite $\mathcal{O}$ d'une part et à W d'autre part, si bien que $g \cdot x$ est adhérent à $\mathcal{O}$. □

3.2. Topologie quotient

3.2.1. Définition. Soit H un sous-groupe d'un groupe topologique G. On définit la *topologie quotient* sur G/H comme la topologie dont les ouverts sont les parties de G/H dont l'image réciproque par la projection canonique $\pi : G \to G/H$ sont ouvertes : pour $\overline{S}$ partie de G/H, $\overline{S}$ est ouvert si $\pi^{-1}(\overline{S})$ est ouverte dans G.

La *saturation* d'une partie Y de G est $\pi^{-1}\big(\pi(Y)\big)$. On dit que Y *saturée* si l'on a : $Y = \pi^{-1}\big(\pi(Y)\big)$. Les ouverts de la forme $\pi^{-1}(\overline{S})$, où $\overline{S}$ parcourt les ouverts de G/H, sont les ouverts saturés de G.

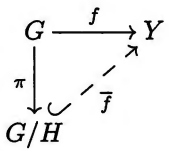
3.2.2. Exercice. *Vérifier que l'ensemble des parties $\overline{S}$ de G/H telles que $\pi^{-1}(\overline{S})$ est ouvert dans G est bien une topologie.*

3.2.3. Lemme. *Soit G un groupe topologique et H un sous-groupe de G. Soit $\pi : G \to G/H$ la projection canonique.*

(i) *L'application π est continue ; plus précisément, la topologie quotient de G/H est la topologie la plus fine pour laquelle π est continue.*

(ii) *L'application π est ouverte.*

(iii) *Soit Y un espace topologique et $f : G \to Y$ une application continue et constante sur les classes modulo H. Alors, l'application $\overline{f}$ qui fait commuter le diagramme suivant est continue :*



Démonstration. (i) Par définition, si $\overline{S}$ est un ouvert de G/H, son image réciproque par π est un ouvert de G, d'où la continuité de π pour la topologie quotient.

Inversement, fixons une topologie pour laquelle π est continue et soit $\overline{S}$ un ouvert de cette topologie. Alors, comme π est continue, $\pi^{-1}(\overline{S})$ est un ouvert de G dont $\overline{S}$ est aussi un ouvert de la topologie quotient.

(ii) Soit V un ouvert de G, on a : $\pi(V) = \pi\left(\bigcup_{h\in H} Vh\right)$. Mais pour tout h de H, Vh est l'image de V par la translation $g \mapsto gh$, qui est un homéomorphisme : donc Vh est un ouvert, si bien que la réunion de Vh aussi. Mais alors, on remarque que : $\bigcup_{h\in H} Vh = \pi^{-1}\big(\pi(V)\big)$. Par suite, $\pi(V)$ est ouvert dans G/H.

(iii) Soit W un ouvert de Y. On veut montrer que $\overline{f}^{-1}(W)$ est ouvert. Pour cela, il suffit de montrer que $V = \pi^{-1}\big(\overline{f}^{-1}(W)\big)$ est ouvert : cela provient du fait que l'on a : $V = (\overline{f}\circ\pi)^{-1}(W) = f^{-1}(W)$ et que f est continue. □

On pourra trouver en annexe D la définition et quelques propriétés d'une topologie naturelle sur un espace quotient X/G, c'est-à-dire sur l'ensemble des orbites pour l'action continue d'un groupe topologique G sur un espace topologique X.

3.2.4. Remarque. Avec la topologie qui convient, l'application $\overline{\alpha_A}$ (du théorème I-3.2) est continue. Effectivement, pour S ouvert dans $G\cdot x$, on a : $\overline{\alpha_x}^{-1}(S) = \pi\big(\alpha_x^{-1}(S)\big)$, qui est un ouvert de la topologie quotient, car $\alpha_x : g \mapsto g\cdot x$ est continue.

3.2.5. Proposition. *Soit G un groupe topologique et soit H un sous-groupe de G. Alors :*

(i) *si H est ouvert, alors H est fermé.*

Si, de plus, G est séparé, alors :

(ii) *le groupe H est fermé si et seulement si G/H est séparé ;*

(iii) *si G/H est séparé, alors l'image d'un compact par π est un compact.*

Démonstration. (i) Supposons H ouvert. Son complémentaire est :

$$G\setminus H = \bigcup_{g\in G\setminus H} gH.$$

Or, chacune des parties gH est ouverte, car c'est l'image réciproque $\tau_{g^{-1}}^{-1}(H)$ de H par l'application continue $\tau_{g^{-1}} : k \mapsto g^{-1}k$. Par suite, $G\setminus H$ est un ouvert et H est un fermé.

(ii) Supposons H fermé. Soit g et g' deux éléments de G tels que $\pi(g) \neq \pi(g')$. On veut séparer $\pi(g)$ et $\pi(g')$. Par hypothèse, $g^{-1}g'$ n'appartient pas à H. Considérons l'application continue

$$\begin{array}{rcl} f : G\times G & \longrightarrow & G \\ (k_1,k_2) & \longmapsto & k_1 g^{-1} g' k_2^{-1}. \end{array}$$

Puisque $g^{-1}g'$ n'appartient pas à H et que H est fermé, $U = G\setminus H$ est un ouvert qui contient $g^{-1}g'$. Comme on a : $f(e,e) = g^{-1}g'$, l'image réciproque $f^{-1}(U)$ est un voisinage de (e,e). Par définition de la topologie produit,

il existe deux ouverts V_1 et V_2 de G tels que l'on ait : $(e,e) \in V_1 \times V_2 \subset f^{-1}(U)$. Soit $W = (V_1 \cap V_2) \cap V_1 \cap V_2)^{-1}$, c'est un ouvert de G contenant e et vérifiant $W = W^{-1}$. Il vient : $W \times W \subset f^{-1}(U)$. Autrement dit, on a : $f(W \times W) \cap H = \emptyset$, soit encore : $Wg^{-1}g'W \cap H = \emptyset$ ou, enfin : $g'WH \cap gWH = \emptyset$. Remarquer que $g'WH$, étant la réunion des ouverts $g'Wh$ lorsque h parcourt H, est ouvert ; de même pour gWH.
Mais alors, $\pi(g'WH)$ et $\pi(gW^{-1}H)$ sont des ouverts de G/H (car la surjection canonique est ouverte), ils contiennent respectivement $\pi(g')$ et $\pi(g)$ et ils sont disjoints. Il en résulte que G/H est séparé.
Inversement, si G/H est séparé, alors le singleton $\{\pi(e)\}$ est fermé, donc son image réciproque l'est aussi ; or c'est : $\pi^{-1}\big\{\big(\pi(e)\big)\big\} = H$.
(iii) Toute application continue entre espaces séparés quelconques envoie un compact sur un compact. Voir, par exemple, [35], chapitre 4. □

3.2.6. Proposition. *Soit G un groupe topologique séparé et H un sous-groupe de G. Les assertions suivantes sont équivalentes :*

(i) *le sous-groupe H est ouvert : $H = \overset{\circ}{H}$;*
(ii) *l'intérieur de H contient le neutre : $e \in \overset{\circ}{H}$;*
(iii) *l'espace topologique G/H est discret.*

Démonstration. Le sous-groupe H contient e : s'il est ouvert, e appartient à l'intérieur de H. Inversement, supposons que l'élement e appartient à l'intérieur $\overset{\circ}{H}$ de H. Soit h un élément de H. Alors $h\overset{\circ}{H}$ est un ouvert de G, contenu dans H, qui contient h, si bien que H est ouvert.
Supposons que le sous-groupe H est ouvert ; alors, il est aussi fermé et π est ouverte. On a : $\{\pi(e)\} = \pi(H)$. Ce singleton est ouvert en tant qu'image de l'ouvert H par l'application ouverte π et fermé en tant que complémentaire de l'image de l'ouvert $G \setminus H$ par π. Par principe de translation, tous les singletons sont ouverts et fermés et la topologie de G/H est discrète.
Inversement, si la topologie quotient sur G/H est discrète, alors H est ouvert en tant qu'image réciproque de la partie $\{\pi(e)\}$ par π. □

3.2.7. Exemple. Voici une illustration simple de cette situation : considérons les groupes orthogonal et spécial orthogonal $G = \mathrm{O}(n)$ et $H = \mathrm{SO}(n)$ (ici, le corps peut être $\mathbb{R}$ ou $\mathbb{C}$). Le sous-groupe H est ouvert dans G comme complémentaire du fermé $\{g \in G, \det g = -1\}$; d'autre part, le déterminant permet d'identifier le quotient G/H à $\{-1,1\}$, un espace discret.

3.2.8. Mise en garde. Le fait qu'un sous-groupe H est ouvert si et seulement si il contient un voisinage de l'élément neutre, que l'on a utilisé ci-dessous, est une application très simple du principe de translation. Mais il s'avère utile dans bien des situations, surtout quand on le couple au théorème de l'application ouverte dans des problèmes de surjectivité, voir le chapitre IX p. 267.

3.3. Théorème d'homéomorphisme pour les groupes compacts

Voici une situation simple et fréquente où l'on peut affirmer qu'une orbite $G \cdot x$ est homéomorphe au quotient G/G_x.

3.3.1. Proposition. *Soit $\alpha : G \times X \to X$ une action continue d'un groupe topologique compact G agissant sur un espace topologique X. Alors, pour tout x de X, l'application $\overline{\alpha_x} : G/G_x \to G \cdot x$, $gG_x \mapsto g \cdot x$, est un homéomorphisme.*

Démonstration. Soit x un élément de X. Rappelons que $\overline{\alpha_x}$ est continue, voir remarque 3.2.4.

Puisque G est compact et que π est continue, G/G_x est compact (puisque séparé par la proposition 3.2.5). Ainsi, tout fermé de G/G_x est compact, si bien que son image par $\overline{\alpha_x}$ est compacte ; par suite, $\overline{\alpha_x}$ est fermée, et comme elle est bijective, elle est ouverte.

Au bilan, puisque $\overline{\alpha_x}$ est bijective, continue et ouverte, c'est un homéomorphisme. □

3.4. Théorème d'homéomorphisme

Pour obtenir des conditions suffisantes moins restrictives permettant d'assurer que $G \cdot x$ soit homéomorphe à G/G_x, il faut introduire des notions plus délicates.

3.4.1. Définition. Soit X un espace topologique. On dit que X est *localement compact* si tout point de X possède un voisinage compact.

On trouvera en annexe C tout un kit de survie en milieu localement compact.

3.4.2. Définition. Soit X un espace topologique. On dit que X est *dénombrable à l'infini* si X peut s'écrire comme une réunion dénombrable de compacts.

3.4.3. Théorème (Théorème d'homéomorphisme). *Soit $\alpha : G \times X \to X$ une action continue et transitive d'un groupe topologique G sur un espace topologique localement compact X. On suppose que G est localement compact et dénombrable à l'infini.*

Alors, l'espace X est homéomorphe au quotient de G par le stabilisateur d'un quelconque élément x de X. Plus précisément, l'application continue

$$\overline{\alpha_x} : G/G_x \to G \cdot x = X, \quad gG_x \mapsto g \cdot x$$

est un homéomorphisme.

3.4.4. Remarque. Tout d'abord, si $\mathbb{K}$ est le corps des réels ou des complexes, $\mathscr{M}_n(\mathbb{K})$ est localement compact et dénombrable à l'infini, car les boules fermées sont compactes en dimension finie. Les hypothèses du théorème s'appliquent donc à $\mathrm{GL}_n(\mathbb{K})$ et à tous ses sous-groupes (ouverts ou) fermés.

En effet, $\mathrm{GL}_n(\mathbb{K})$ est un ouvert de l'espace vectoriel $\mathscr{M}_n(\mathbb{K})$: les ouverts pour la topologie de $\mathrm{GL}_n(\mathbb{K})$ sont donc aussi des ouverts pour celle de $\mathscr{M}_n(\mathbb{K})$. Donc $\mathrm{GL}_n(\mathbb{K})$ est localement compact et dénombrable à l'infini. De plus, un sous-groupe fermé de $\mathrm{GL}_n(\mathbb{K})$ l'est également.

Démonstration. Soit x un élément de X. Comme $\overline{\alpha_x}$ est bijective et continue, il suffit de montrer que c'est une application ouverte, ce qui sera le cas si $\alpha_x : G \to X$, $g \mapsto g \cdot x$ l'est. Il s'agit donc de montrer que si U est un ouvert de G, alors $\alpha_x(U) = U \cdot x$ est un ouvert de X.

Soit g un élément de U, montrons que $g \cdot x$ est intérieur à $U \cdot x$; pour cela, par principe de translation, il suffit de montrer que x est intérieur à $g^{-1} \cdot (U \cdot x) = (g^{-1}U) \cdot x$.

Puisque $g^{-1}U$ est un voisinage du neutre e, il existe un voisinage compact V de e tel que $e \in V \subset g^{-1}U$. Par continuité du produit de G, il existe un voisinage compact W de e tel que $W^2 \subset V$. Quitte à remplacer W par l'intersection $W \cap W^{-1}$, on peut supposer que $W = W^{-1}$.

Comme W est un voisinage de e, hW est un voisinage de h pour tout h. En particulier, pour tout compact $K \subset G$, la famille $(k \overset{\circ}{W})_{k \in K}$ est un recouvrement ouvert de K, et l'on peut donc en extraire un sous-recouvrement fini $(k_i \overset{\circ}{W})_{i=1,\dots,N}$. Comme G est dénombrable à l'infini, on peut trouver une suite d'éléments $(g_i)_{i \in \mathbb{N}}$ tels que G soit recouvert par $(g_i W)_{i \in \mathbb{N}}$.

Puisque l'action de G est transitive, on peut écrire :

$$X = G \cdot x = \Big(\bigcup_{i \in \mathbb{N}} g_i W \Big) \cdot x = \bigcup_{i \in \mathbb{N}} g_i W \cdot x.$$

Puisque X est localement compact, c'est un espace de Baire (voir, en annexe, le théorème C.7). Puisque l'intérieur de X est non vide et réunion dénombrable de fermés (chaque $g_i W \cdot x$ est compact comme image d'un compact par α_x), il existe un fermé d'intérieur non vide, disons $g_{i_0} W \cdot x$.

Soit $g_{i_0} w \cdot x$ un point intérieur à $g_{i_0} W \cdot x$. Alors, par translation, x est intérieur à

$$w^{-1} g_{i_0}^{-1} \cdot (g_{i_0} W \cdot x) = w^{-1} W \cdot x \subset (W^2) \cdot x \subset (g^{-1}U) \cdot x,$$

ce dont il fallait s'assurer. □

3.4.5. Définition. Si G est un groupe topologique localement compact, dénombrable à l'infini qui agit transitivement sur un espace topologique X, on dit que X est un *espace homogène* (sous l'action de G).

3.4.6. Exemple. Sous les hypothèses du théorème, si H est un sous-groupe fermé de G, G/H est un espace homogène pour l'action par translation à gauche. Le théorème d'homéomorphisme montre que tous les espaces homogènes sont homéomorphes à un espace de cette forme.

4. Applications du théorème d'homéomorphisme

Comme tout bon théorème, le théorème d'homéomorphisme a plusieurs types d'applications :

1. comprendre des aspects topologiques du groupe G grâce à son action sur X, par l'étude de X et de G_x ; (c'est la philosophie de la théorie des représentations...)
2. comprendre des aspects topologiques de X à partir de l'action de G ;
3. définir une topologie « légitime » sur un ensemble X à partir de l'action d'un groupe topologique G ; c'est une première étape pour « géométriser » l'ensemble X.

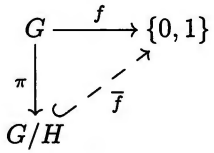
Les trois situations seront étudiées dans les sections qui suivent.

4.1. Étude de la connexité d'un groupe *via* une action

Pour commencer, voici une propriété générale de la topologie quotient qui sera utile pour les cas concrets.

4.1.1. Proposition. *Soit G un groupe topologique et H un sous-groupe de G. Si H et G/H sont connexes, alors G est connexe.*

Démonstration. Pour montrer que G est connexe, considérons une application continue $f : G \to \{0,1\}$ et montrons que f est constante. Tout d'abord, f est continue sur H donc, par connexité de H, constante sur H. De même, toutes les classes gH, $g \in G$ sont homéomorphes à H donc connexes, si bien que f est constante sur chaque classe. On peut passer au quotient : il existe $\overline{f} : G/H \to \{0,1\}$ faisant commuter le diagramme :



D'après le lemme 3.2.3 (iii), $\overline{f}$ est continue. Par connexité de G/H, elle est donc constante. Par surjectivité de π, f est également constante. Ainsi, G est connexe. □

Suivant [25], chapitre 2, exercice 4, on peut démontrer, dans le même esprit l'énoncé ci-après.

4.1.2. Proposition. *Soit G un groupe topologique et H un sous-groupe de G. Si H et G/H sont compacts, alors G est compact.*

Nous avons déjà prouvé la connexité de $\mathrm{GL}_n(\mathbb{C})$ dans la proposition 1.4. Nous allons en donner une autre preuve fondée sur le théorème d'homéomorphisme pour illustrer une méthode générale.

4.1.3. Proposition. *Pour tout entier n, le groupe $\mathrm{GL}_n(\mathbb{C})$ est connexe.*

Démonstration. On procède par récurrence sur n. Pour $n = 1$, on a : $\mathrm{GL}_1(\mathbb{C}) = \mathbb{C}^*$, qui est bien un groupe connexe.

Soit n entier supérieur à 2, supposons que $\mathrm{GL}_{n-1}(\mathbb{C})$ soit connexe. L'action naturelle de $\mathrm{GL}_n(\mathbb{C})$ sur $\mathbb{C}^n$, le produit d'une matrice par un vecteur, se restreint en une action sur $\mathbb{C}^n \setminus \{0\}$. Cette action est transitive. En effet, si v et w sont deux vecteurs non nuls, on complète les familles (v) et (w) en deux bases $\mathbf{v} = (v, v_2, \ldots, v_n)$ et $\mathbf{w} = (w, w_2, \ldots, w_n)$ de $\mathbb{C}^n$. Soit g la matrice de changement de base de $\mathbf{v}$ à $\mathbf{w}$, on a : $gv = w$.

Le stabilisateur d'un élément, par exemple le premier vecteur de la base naturelle de $\mathbb{C}^n$, est décrit par blocs par :

$$H = \left\{ \begin{pmatrix} 1 & \ell \\ 0 & h \end{pmatrix} \ , \ \ell \in \mathscr{M}_{1,n-1}(\mathbb{C}),\ h \in \mathrm{GL}_{n-1}(\mathbb{C}) \right\},$$

$$\text{noté : } H = \begin{pmatrix} 1 & \mathbb{C}^{n-1} \\ 0 & \mathrm{GL}_{n-1}(\mathbb{C}) \end{pmatrix}.$$

On voit qu'en tant qu'espace topologique, H est homéomorphe au produit $\mathbb{C}^{n-1} \times \mathrm{GL}_{n-1}(\mathbb{C})$. Par hypothèse de récurrence, H est connexe. De plus, d'après le théorème d'homéomorphisme 3.4.3 (qu'on peut appliquer d'après la remarque 3.4.4), le quotient $\mathrm{GL}_n(\mathbb{C})/H$ est homéomorphe à $\mathbb{C}^n \setminus \{0\}$, qui est connexe. D'après la proposition 4.1.1, $\mathrm{GL}_n(\mathbb{C})$ est connexe, ce qui termine la récurrence. □

Voici un mot du vocabulaire standard.

4.1.4. Définition. On appelle *composante neutre* d'un groupe topologique la composante connexe qui contient l'élément neutre.

4.1.5. Proposition. *Soit n un entier naturel non nul. Le groupe*

$$\mathrm{GL}_n^+(\mathbb{R}) = \{g \in \mathrm{GL}_n(\mathbb{R}) \ , \ \det g > 0\}$$

est connexe. C'est la composante neutre de $\mathrm{GL}_n(\mathbb{R})$.

Démonstration. La même preuve par récurrence que ci-dessus fonctionne : $\mathrm{GL}_n^+(\mathbb{R})$ agit transitivement sur $\mathbb{R}^n \setminus \{0\}$, qui est connexe pour $n \geqslant 2$, et les stabilisateurs sont connexes.

Soit g_0 la matrice diagonale dont tous les coefficients diagonaux valent 1, sauf le dernier qui vaut -1. Par multiplicativité du déterminant, $g_0\,\mathrm{GL}_n^+(\mathbb{R})$ est l'ensemble $\mathrm{GL}_n^-(\mathbb{R})$ des matrices de déterminant négatif. Puisque la translation $L_{g_0} : \mathrm{GL}_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R})$, $g \mapsto g_0 g$ est un homéomorphisme, $\mathrm{GL}_n^-(\mathbb{R})$ est connexe. En revanche, la réunion $\mathrm{GL}_n(\mathbb{R})$ n'est pas connexe car l'application $g \mapsto \det(g)/|\det(g)|$ est une surjection continue sur $\{-1, 1\}$. On peut conclure. □

4.1.6. Proposition. *Soit n un entier naturel non nul. Le groupe orthogonal $\mathrm{SO}(n, \mathbb{R})$ est compact et connexe. Le groupe $\mathrm{O}(n, \mathbb{R})$ est également compact, mais non connexe ; sa composante neutre est $\mathrm{SO}(n, \mathbb{R})$.*

Démonstration. On procède par récurrence. La propriété est évidente pour $n = 1$. Pour $n \geqslant 2$, le groupe $\mathrm{SO}(n, \mathbb{R})$ agit transitivement sur la sphère unité S, qui est compacte et connexe. Avec le théorème de la base incomplète et le procédé de Gram-Schmidt, on prouve que l'action est transitive. Le stabilisateur d'un point est homéomorphe à SO_{n-1}. On peut conclure avec les propositions 4.1.1 et 4.1.2.

Pour la compacité, on peut préférer la preuve plus simple suivante : $\mathrm{O}(n, \mathbb{R})$ est constitué d'éléments de norme 1 pour la norme subordonnée $|||\cdot|||_2$; il est donc borné. De plus, la condition ${}^tAA = \mathrm{I}_n$ définit un fermé, donc $\mathrm{O}(n, \mathbb{R})$ est fermé borné, donc compact. Il en est de même de $\mathrm{SO}(n, \mathbb{R})$ qui est l'intersection de $\mathrm{O}(n, \mathbb{R})$ avec le fermé $\mathrm{SL}_n(\mathbb{R})$.

La condition ${}^tAA = \mathrm{I}_n$ implique que le déterminant envoie $\mathrm{O}(n, \mathbb{R})$ sur $\{1, -1\}$ et ce, de façon clairement surjective : prendre la matrice g_0 ci-dessus, *i.e.* la diagonale $(-1, 1, \cdots, 1)$. Donc, $\mathrm{O}(n, \mathbb{R})$ n'est pas connexe. Plus précisément, $\mathrm{O}(n, \mathbb{R})$ est la réunion non connexe des deux parties connexes $\mathrm{SO}(n, \mathbb{R})$ et $g_0\,\mathrm{SO}(n, \mathbb{R})$, si bien que $\mathrm{SO}(n, \mathbb{R})$ est la composante neutre de $\mathrm{O}(n, \mathbb{R})$. □

4.1.7. Exercice. *Prouver que les groupes unitaires sont compacts.*

4.2. Connexité des classes d'équivalence de matrices complexes

Prouvons la connexité des classes d'équivalence de matrices complexes grâce à l'action de groupes linéaires. Soit m, n entiers. Rappelons l'action de $\mathrm{GL}_m(\mathbb{C}) \times \mathrm{GL}_n(\mathbb{C})$ sur $\mathscr{M}_{m,n}(\mathbb{C})$ par :

$$\forall (P, Q) \in \mathrm{GL}_m(\mathbb{C}) \times \mathrm{GL}_n(\mathbb{C}),\ \forall A \in \mathscr{M}_{m,n}(\mathbb{C}), (P, Q) \cdot A = PAQ^{-1}.$$

Les orbites $\mathscr{O}_r$ sont les classes d'équivalence de matrices, elles sont paramétrées par le rang $r \in \{0, \ldots, \min(m, n)\}$.

4.3. Structure topologique sur les grassmanniennes

Dans cette section, on définit les grassmanniennes qui seront souvent utilisée par la suite dans des contextes différents. On pourra donc supposer que $\mathbb{K}$ sauf mention du contraire est un corps quelconque.

4.3.1. Définition. Soit m et n deux entiers naturels tels que $0 \leqslant m \leqslant n$. On appelle *grassmannienne des m-plans de $\mathbb{K}^n$* et l'on note $\mathrm{Gr}_{m,n}(\mathbb{K})$ ou $\mathrm{Gr}_m(\mathbb{K}^n)$ l'ensemble des sous-espaces vectoriels de dimension m de $\mathbb{K}^n$:

$$\mathrm{Gr}_{m,n}(\mathbb{K}) = \{F \subset \mathbb{K}^n \ , \ \ F \text{ sous-e.v.}, \ \dim F = m\}.$$

4.3.2. Remarque. Lorsque $m = 0$ ou $m = n$, la grassmannienne est une singleton : pas passionnant.

Lorsque $m = 1$, la grassmannienne $\mathrm{Gr}_{1,n}(\mathbb{K})$ des droites de $\mathbb{K}^n$ est l'espace projectif $\mathbb{P}^{n-1}(\mathbb{K})$ de dimension $n-1$, que l'on étudiera plus en détail plus loin. Voir aussi l'exercice F.23.

Lorsque $m = n-1$, la grassmannienne $\mathrm{Gr}_{n-1}(\mathbb{K}^n)$ des hyperplans de $\mathbb{K}^n$ s'identifie à l'espace projectif $\mathrm{Gr}_1\big((\mathbb{K}^n)^*\big)$ du dual de $\mathbb{K}^n$: comment ?

L'action naturelle de $\mathrm{GL}_n(\mathbb{K})$ sur $\mathbb{K}^n$ en induit une sur $\mathrm{Gr}_{m,n}(\mathbb{K})$. Grâce au théorème de la base incomplète, on prouve que cette action est transitive. Soit en effet F_0 le sous-espace engendré par les m premiers vecteurs de la base naturelle $\mathbf{e}$ de $\mathbb{K}^n$; pour F élément de $\mathrm{Gr}_{m,n}(\mathbb{K})$, soit $(v_1, \ldots, v_m)$ une base de F ; on la complète en une base $\mathbf{v} = (v_1, \ldots, v_n)$ de $\mathbb{K}^n$; la matrice de passage g de la base $\mathbf{e}$ à la base $\mathbf{v}$ envoie F_0 sur F.

4.3.3. Remarque. L'application orbitale

$$\alpha_{F_0} : \mathrm{GL}_n(\mathbb{K}) \longrightarrow \mathrm{Gr}_{m,n}(\mathbb{K}), \quad g \longmapsto g \cdot F_0$$

envoie une matrice g sur le sous-espace vectoriel engendré par ses m premières colonnes. On vient de montrer qu'elle est surjective.

Le stabilisateur de F_0 est décrit par blocs par :

$$P_{m,n-m} = \begin{pmatrix} \mathrm{GL}_m(\mathbb{K}) & \mathscr{M}_{m,n-m}(\mathbb{K}) \\ 0 & \mathrm{GL}_{n-m}(\mathbb{K}) \end{pmatrix}.$$

Le groupe $P_{m,n-m}$ est un exemple de ce que l'on appelle *groupe parabolique* : il est isomorphe à un produit semi-direct (voir ci-dessous) :

$$P_{m,n-m} \simeq \mathscr{M}_{m,m-n}(\mathbb{K}) \rtimes \big(\mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_{n-m}(\mathbb{K})\big).$$

Ainsi, on peut définir une topologie sur $\mathrm{Gr}_{m,n}(\mathbb{K})$ par transport de structure en utilisant la topologie quotient de $\mathrm{GL}_n(\mathbb{K})/P_{m,n-m}$ et la bijection provenant de l'application orbitale $\mathrm{GL}_n(\mathbb{K})/P_{m,n-m} \to \mathrm{Gr}_{m,n}(\mathbb{K})$, $g \mapsto gF_0$.

4.3.4. Définition (Topologie sur les grassmanniennes). Soit $\mathbb{K}$ le corps des réels ou des complexes et soit m et n deux entiers tels que $0 \leqslant m \leqslant n$. On munit $\mathrm{GL}_n(\mathbb{K})/P_{m,n-m}$ de la topologie quotient. On munit $\mathrm{Gr}_{m,n}(\mathbb{K})$ de la topologie induite par transport de structure (qui fait de la bijection un homéomorphisme)

4.3.5. Exercice. *Montrer, à l'aide du principe de conjugaison, que cette topologie ne dépend pas du choix du point de base F_0.*

Continuons l'étude de l'action de $\mathrm{GL}_n(\mathbb{K})$ sur $\mathrm{Gr}_{m,n}(\mathbb{K})$. Celle-ci n'est pas fidèle puisque les homothéties agissent trivialement. En fait, et il s'agit là d'un grand classique, parfois appelé lemme de Schur.

4.3.6. Proposition. *Le noyau de l'action de $\mathrm{GL}_n(\mathbb{K})$ sur l'espace projectif $\mathbb{P}^{n-1}(\mathbb{K})$ est réduit aux homothéties.*

Démonstration. Si g dans $\mathrm{GL}_n(\mathbb{K})$ laisse fixe toute droite alors pour tout couple de vecteurs (v, w) non nuls, on a $g(v) = \lambda v$, $g(w) = \mu v$, $\lambda, \mu \in \mathbb{K}$. Si les vecteurs v et w sont colinéaires, alors le fait qu'ils soient non nuls implique que $\lambda = \mu$. S'il ne le sont pas, ils forment une partie libre et de plus, on a :

$$\lambda v + \mu w = g(v) + g(w) = g(v + w) = \nu(v + w) = \nu v + \nu w.$$

Par identification, il vient donc : $\lambda = \nu = \mu$. Ainsi, g est bien une homothétie. □

Ce résultat se généralise à toute grassmannienne. Soit m fixé dans $[\![1, n-1]\!]$ et g dans $\mathrm{GL}_n(\mathbb{K})$ stabilisant tout sous-espace de dimension m, alors g est une homothétie. Ceci provient tout simplement du théorème de la base incomplète ! Soit D une droite quelconque et e un vecteur directeur de cette droite. On peut compléter e en une base $(e_1, e_2, \ldots, e_n)$ de $\mathbb{K}^n$, avec $e = e_1$. Comme m est dans $[\![1, n-1]\!]$, l'intersection des parties à m éléments de $\{e_1, e_2, \ldots, e_n\}$ contenant e_1 est exactement $\{e_1\}$. Comme g fixe tous les sous-espaces de dimension m qu'elles engendrent, g doit fixer D. Donc, g fixe toute droite et, ainsi, g est une homothétie.
Il est temps de faire le point sur les résultats obtenus.

4.3.7. Définition. Soit n un entier. On appelle *groupe projectif* (linéaire) et l'on note $\mathrm{PGL}_n(\mathbb{K})$ le quotient de $\mathrm{GL}_n(\mathbb{K})$ par le sous-groupe distingué des homothéties.

La définition est motivée par le corollaire suivant.

4.3.8. Corollaire. *L'action naturelle (transitive) de $\mathrm{GL}_n(\mathbb{K})$ sur $\mathbb{K}^n \backslash \{0\}$ définit par passage au quotient une action fidèle et transitive de $\mathrm{PGL}_n(\mathbb{K})$ sur $\mathbb{P}^{n-1}(\mathbb{K})$.*

4.3.9. Proposition. *Le groupe* $\mathrm{GL}_n(\mathbb{K})$ *agit transitivement sur la grassmannienne* $\mathrm{Gr}_{m,n}(\mathbb{K})$ *pour tout* m *de* $[\![0,n]\!]$. *Le stabilisateur d'un élément étant isomorphe au sous-groupe parabolique* $P_{m,n-m}$.
Si $m \notin \{0,n\}$, *alors le noyau de l'action est réduit aux homothéties. Autrement dit,* $\mathrm{PGL}_n(\mathbb{K})$ *agit fidèlement sur* $\mathrm{Gr}_{m,n}(\mathbb{K})$.

Jusqu'ici, au fond, nous n'avions pas besoin du théorème d'homéomorphisme. Celui-ci ne sert que de ligne directrice pour définir la topologie sur la grassmannienne à partir de celle de l'espace homogène. En revanche, une réciproque simplissime du théorème d'homéomorphisme va nous être utile.

4.3.10. Proposition. *Soit* G *un groupe topologique agissant transitivement sur un espace topologique* X *et soit* G_x *le stabilisateur d'un élément* x *de* X. *On suppose que l'application orbitale* $g \mapsto g \cdot x$ *fournit un homéomorphisme* ϕ_x *de* G/G_x *dans* X. *Alors, l'action de* G *sur* X *est continue.*

Démonstration. Il faut montrer avec les hypothèses que l'application $\alpha : G \times X \to X$, $(g,x) \mapsto g \cdot x$, est continue. L'homéomorphisme ϕ_x permet de construire le diagramme

$$\begin{array}{ccc} G \times X & \xrightarrow{\alpha} & X \\ {\scriptstyle \mathrm{Id} \times \alpha_x^{-1}} \downarrow & & \uparrow {\scriptstyle \phi_x} \\ G \times G/G_x & \xrightarrow{\psi} & G/G_x \end{array}$$

$$(h, gG_x) \longrightarrow hgG_x.$$

Comme α est une action, on a pour tout $y = g \cdot x$ de X et pour tout h de G les relations : $\alpha(h,y) = h \cdot y = (hg) \cdot x = \phi_x\big(\psi(h, gG_x)\big)$, ce qui implique que le diagramme est commutatif. Comme ψ est un passage au quotient de la multiplication de G, elle est continue par construction de la topologie de G/H et donc α est continue. □

Supposons ici que $\mathbb{K} = \mathbb{R}$...

4.3.11. Proposition. *Soit* m, n *deux entiers tels que* $1 \leqslant m < n$. *La grassmannienne réelle* $\mathrm{Gr}_{m,n}(\mathbb{R})$ *est compacte.*

Démonstration. L'action de $\mathrm{GL}_n(\mathbb{R})$ sur $\mathrm{Gr}_{m,n}(\mathbb{K})$ continue par la proposition précédente. Elle se restreint en une action continue du groupe orthogonal $\mathrm{O}(n,\mathbb{R})$. Cette action est toujours transitive : étant donné un sous-espace F de dimension m, on en choisit une base $(v_1,\dots,v_m)$, on la complète en une base $(v_1,\dots,v_n)$ de $\mathbb{R}^n$, on applique le procédé d'orthogonalisation de Gram-Schmidt pour obtenir une base orthonormée $(w_1,\dots,w_n)$ de $\mathbb{R}^n$ telle que $\mathbf{w} = (w_1,\dots,w_m)$ est une base orthonormée de F. La matrice de passage de la base canonique à $\mathbf{w}$ est orthogonale et elle envoie le sous-espace F_0 utilisé ci-dessus sur F.

Mais alors, par le théorème d'isomorphisme, la grassmannienne est homéomorphe à un quotient du groupe compact $\mathrm{O}(n,\mathbb{R})$: il en résulte que $\mathrm{Gr}_{m,n}(\mathbb{R})$ est compacte. □

4.3.12. Remarque. On peut montrer que les grassmanniennes complexes sont aussi compactes en utilisant le groupe unitaire à la place du groupe orthogonal. Voir l'exercice F.20.

5. Produits semi-directs topologiques

La motivation principale de ce paragraphe, c'est qu'un grand nombre des groupes qui apparaissent naturellement dans des situations algébriques ou géométriques, en particulier dans la suite de ce livre, sont des produits semi-directs : groupe des homothéties-translations, groupe affine du plan, groupe des isométries affine du plan, ou, dans l'exemple ci-dessus, stabilisateur d'un sous-espace dans un groupe linéaire, etc.

Une motivation plus interne à la théorie des groupes est la question suivante : supposons connaître un groupe (topologique) G seulement par un sous-groupe distingué K (comme *kernel*) et le quotient correspondant $H = G/K$, peut-on reconstruire G ? Les produits semi-directs donnent une réponse possible mais, à certains égards[5], elle est triviale.

5.1. Définition du produit semi-direct topologique

5.1.1. Définition. Soit H un groupe topologique agissant continûment par automorphismes sur un groupe topologique K. Notons le morphisme associé à l'action :

$$\begin{array}{rcl} \varphi : H & \longrightarrow & \mathrm{Aut}\, K \\ h & \longmapsto & \varphi_h, \end{array}$$

où $\mathrm{Aut}\, K$ est le groupe des automorphismes de K.

On définit le *produit semi-direct topologique*, noté $K \rtimes_\varphi H$ ou simplement $K \rtimes H$, ainsi :

- comme espace topologique, c'est le produit cartésien $K \times H$, muni de la topologie produit,
- on le munit du produit défini par :

$$(*) \qquad \forall (k,h),(k',h') \in K \rtimes_\varphi H, \quad (k,h)(k',h') = \big(k\varphi_h(k'), hh'\big).$$

5.1.2. Mise en garde. Il peut arriver que l'on note $h \cdot k = \varphi_h(k)$ pour l'action d'un élément h de H sur un élément k de K. Attention, alors, à ne pas confondre l'action $h \cdot k$ d'un élément h de H sur un élément k de K et la loi de groupe dans le produit semi-direct.

[5] Pour épater la galerie, on pourrait invoquer la cohomologie !

5.1.3. Lemme.

(i) *Avec les notations précédentes, $K \rtimes_\varphi H$ est un groupe topologique.*

(ii) *Ce groupe possède naturellement un sous-groupe fermé isomorphe à H, un sous-groupe fermé distingué isomorphe à K, et le quotient par ce dernier sous-groupe est isomorphe à H.*

Démonstration. (i) Tout d'abord, le produit défini par la relation $(*)$ est bien une loi interne. L'élément (e_K, e_H) de $K \rtimes H$ est neutre pour le produit (où on a noté e_K et e_H les neutres de K et H).
Pour l'associativité, fixons trois éléments $(k,h),(k',h'),(k'',h'') \in K \rtimes H$. On a :

$$\begin{aligned}(k,h)\,[(k',h')(k'',h'')] &= (k,h)\big(k'\varphi_{h'}(k''), h'h''\big)\\ &= \big(k\varphi_h(k'\varphi_{h'}(k'')), hh'h''\big)\\ &= \big(k\varphi_h(k')\varphi_{hh'}(k''), hh'\,h''\big)\\ &= \big(k\varphi_h(k'), hh'\big)(k'',h'')\\ &= [(k,h)(k',h')]\,(k'',h'').\end{aligned}$$

Nous avons utilisé librement le fait que φ_h est un automorphisme pour tout h et que φ est un morphisme.
Étant donné (k,h) de $K \rtimes H$, on a :

$$(k,h)\big(\varphi_{h^{-1}}(k^{-1}), h^{-1}\big) = (e_K, e_H) = \big(\varphi_{h^{-1}}(k^{-1}), h^{-1}\big)(k,h),$$

si bien que $\iota(k,h) = \big(\varphi_{h^{-1}}(k^{-1}), h^{-1}\big)$ est un inverse de (k,h).
Enfin, la continuité du produit et de l'inversion ι résulte de celle des opérations de K et H. Cela prouve l'assertion (i).
(ii) D'abord, considérons les applications : $K \to K \rtimes H$, $k \mapsto (k, e_H)$ et $H \to K \rtimes H$, $h \mapsto (e_K, h)$. Elles sont continues, injectives, et l'on vérifie avec la loi donnée en $(*)$ que ce sont des morphismes de groupes.
Écrivons alors la formule de conjugaison d'un élément (k, e_H) de $K \times \{e_H\}$ par un élément (e_K, h) de $\{e_K\} \times H$:

$$(**) \qquad (e_K,h)(k,e_H)(e_K,h)^{-1} = \big(\varphi_h(k), h\big)(e_K, h^{-1}) = \big(\varphi_h(k), e_H\big).$$

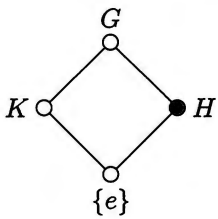
En particulier, $K \times \{e_H\}$ est distingué dans le produit semi-direct.
Enfin, par la relation $(*)$, l'application $K \rtimes H \to H$, $(k,h) \mapsto h$ est un morphisme continu de groupes dont le noyau est $K \times \{e_H\}$. D'où l'assertion (ii). □

5.1.4. Remarque. La formule $(*)$ définissant le produit semi-direct paraît artificielle, mais elle est essentiellement équivalente à la formule $(**)$ qui donne, en simplifiant les notations avec les injections ci-dessus : $hkh^{-1} = \phi_h(k)$ (pour $h \in H$ et $k \in K$). On peut interpréter cette formule ainsi : le produit semi-direct est un groupe où H agit sur K par conjugaison de la façon prescrite par φ. La définition est donc, somme toute, naturelle...

5.1.5. Remarque. Remarquons aussi que l'intersection de $K \times \{e_H\}$ et $\{e_K\} \times H$ est triviale, ce qu'on écrit en simplifiant : $K \cap H = \{(e_K, e_H)\}$. Enfin, pour $(k,h) \in K \rtimes H$: $(k,h) = (k,e_H)(e_K,h)$, d'où notamment : $K \rtimes H = (K \times \{e_H\})(\{e_K\} \times H)$, ce qu'on peut écrire : $K \rtimes H = KH$.

On résume la construction dans les deux icônes ci-contre.

$$G \overset{\text{esp. top.}}{\simeq} K \times H,$$



5.2. Critère de reconnaissance d'un produit semi-direct

La proposition suivante permet de reconnaître des produits semi-directs « dans la nature ». Il peut être éclairant d'en comparer les hypothèses avec les ingrédients de la remarque ci-dessus.

5.2.1. Proposition. *Soit G un groupe topologique et soit H et K deux sous-groupes tels que :*

1. *$K \lhd G$, c'est-à-dire que K est distingué dans G,*
2. *$K \cap H = \{e\}$, c'est-à-dire que l'intersection de H et K est triviale,*
3. *$KH = G$, c'est-à-dire que tout élément de G peut s'écrire comme produit d'un élément de K et d'un élément de H.*

Alors, l'application $\mu : (k,h) \mapsto kh$ est un isomorphisme du produit semi-direct $K \rtimes_\varphi H$ vers G, où l'action de H sur K est induite par la conjugaison dans G :

$$\begin{aligned} \varphi : H &\longrightarrow \mathrm{Aut}(K) \\ h &\longmapsto (\varphi_h : k \mapsto hkh^{-1}). \end{aligned}$$

Si, de plus, G est localement compact et dénombrable à l'infini et si K est fermé dans G, alors l'isomorphisme est aussi un homéomorphisme.

Démonstration. Par construction, l'application $\mu : K \times H \to G$, $(k,h) \mapsto kh$ est continue, surjective par *3.* et injective par *2.* Reste à voir que c'est un morphisme de groupes. Pour $(h,k),(h',k')$ de $H \times K$, on a dans G :

$$\begin{aligned} \mu(k,h)\mu(k',h') &= (kh)(k'h') = (khk'h^{-1})(hh') \\ &= \big(k\varphi_h(k')\big)(hh') = \mu\big((h,k)(h',k')\big), \end{aligned}$$

où le produit $(h,k)(h',k')$ est calculé dans le produit semi-direct $K \rtimes_\varphi H$. À présent, supposons G localement compact et dénombrable à l'infini et K fermé. Considérons l'action à gauche de H sur le quotient G/K : pour h de H et gK de G/K, on définit : $h \cdot (gK) = hgK$. Par l'hypothèse *3.*, cette

action est transitive, c'est-à-dire que $H \cdot eK = G/K$: en effet, comme $G = HK$ (ce qui équivaut à $G = KH$ par passage à l'inverse), si $g \in G$, on peut écrire $g = hk$ avec $(h,k) \in H \times K$ et on a $gK = hK$. D'après l'hypothèse *2.*, le stabilisateur de eK dans H est trivial : $H_{eK} = \{e_H\}$. En identifiant H et $H/\{e_H\}$, on obtient par le théorème 3.4.3 un homéomorphisme $\overline{\alpha_{eH}}$: $H = H/\{e_H\} \to G/K$, $h \mapsto hK$. Notons $\sigma : G/K \to H$ l'homéomorphisme inverse de $\overline{\alpha_{eH}}$. Pour h de H, l'égalité $\overline{\alpha_{eH}}(h) = hK$ donne : $\sigma(hK) = h$.

On veut calculer l'inverse de μ. Soit $g \in G$. On sait qu'il existe un unique couple $(k,h) \in K \times H$ tel que $g = kh$ et alors, on a : $\mu^{-1}(g) = (k,h)$. Mais, on peut écrire : $g = h(h^{-1}kh)$, d'où en projetant dans G/K : $gK = hK$. Par suite, il vient : $\sigma(gK) = \sigma(hK) = h$. Enfin, on a : $k = gh^{-1} = g\sigma(gK)^{-1}$. Ainsi, l'inverse de μ est l'application continue suivante :

$$G \longrightarrow K \rtimes_\varphi H, \quad g \longmapsto \big(g\sigma(gK)^{-1}, \sigma(gK)\big).$$

Cela prouve que μ est un homéomorphisme. □

5.2.2. Mise en garde. Comme on peut le constater sur certains forums spécialisés, cette proposition a induit en erreur nos meilleurs étudiants, en particulier, notre étudiant fétiche, Marc Miroil. En effet, il y a deux situations que l'on doit bien distinguer :

- d'une part, la situation où l'on se donne deux groupes K et H existant de façon indépendante et où l'on ajoute la donnée d'un morphisme de H vers $\operatorname{Aut} K$;
- d'autre part, la situation où H et K sont deux sous-groupes d'un groupe G soumis aux hypothèses de la proposition.

Dans le premier cas, c'est la définition du produit semi-direct qui nous permet de créer un nouveau groupe G. Dans le second cas, le groupe G existe déjà, mais on le réalise comme produit semi-direct ; alors, le morphisme de H vers $\operatorname{Aut} K$ est imposé par la conjugaison dans G.

5.3. Reformulation en termes de suites exactes

5.3.1. Définition. Considérons une suite finie de groupes (topologiques) $(G_i)_{i=0,\dots,\ell}$ reliés par des morphismes (continus) $\varphi_i : G_i \to G_{i+1}$ $(0 \leqslant i < \ell)$. On résume cela dans un diagramme :

$$G_0 \xrightarrow{\varphi_0} G_1 \xrightarrow{\varphi_1} G_2 \xrightarrow{\varphi_2} \cdots \xrightarrow{\varphi_{\ell-2}} G_{\ell-1} \xrightarrow{\varphi_{\ell-1}} G_\ell.$$

On dit que cette suite est *exacte* si, pour tout i compris entre 1 et $\ell - 1$, on a :

$$\operatorname{Im} \varphi_{i-1} = \operatorname{Ker} \varphi_i.$$

La situation la plus courante est celle où on a trois groupes K, G et Γ et des morphismes $\iota : K \to G$ et $\pi : G \to \Gamma$. On peut écrire une suite de

groupes (en notant 1 le groupe trivial) :

$$1 \longrightarrow K \xrightarrow{\iota} G \xrightarrow{\pi} \Gamma \longrightarrow 1.$$

Cette suite est exacte si et seulement si ι et injective, π est surjective et $\operatorname{Im}\iota = \operatorname{Ker}\pi$. On parle alors de *suite exacte courte.* Le morphisme ι identifie K à un sous-groupe distingué de G et, par passage au quotient, π induit un isomorphisme entre G/K et Γ.

5.3.2. Rappel (Sections et rétractions). Au niveau *ensembliste*, c'est-à-dire sans considérer de structures de groupe ou autre, on sait que :
- une application $\iota : K \to G$ est injective si et seulement si elle est inversible à gauche, c'est-à-dire s'il existe une application $\rho : G \to K$ telle que $\rho \circ \iota = \mathrm{Id}_K$ (on parle de « *rétraction ensembliste* ») ;
- une application $\pi : G \to \Gamma$ est surjective si et seulement si elle est inversible à droite, *i.e.* s'il existe une application $\sigma : \Gamma \to G$ telle que $\pi \circ \sigma = \mathrm{Id}_\Gamma$ (on parle de « *section ensembliste* »).

Dès que l'on ajoute de la structure, les choses se compliquent. Dans une suite exacte courte, il n'existe pas nécessairement de section, on va voir que l'existence de sections qui soient des morphismes de groupes est essentiellement équivalente à une situation de produit semi-direct.

5.3.3. Exercice. *Construire la suite exacte de groupes :*

$$1 \longrightarrow \mathbb{Z}/4\mathbb{Z} \xrightarrow{\iota} \mathbb{Z}/8\mathbb{Z} \xrightarrow{\pi} \mathbb{Z}/2\mathbb{Z} \longrightarrow 1.$$

Exhiber une section ensembliste de π (respectivement une rétraction ensembliste de ι), mais prouver qu'il n'en existe pas qui soit un morphisme de groupes.

Étant donné un produit semi-direct $K \rtimes_\varphi H$, on obtient grâce au lemme 5.1.3 une suite exacte courte munie d'une section :

$$1 \longrightarrow K \underset{\rho}{\overset{\iota}{\rightleftarrows}} K \rtimes_\varphi H \underset{\sigma}{\overset{\pi}{\rightleftarrows}} H \longrightarrow 1,$$

où $\pi : (k, h) \mapsto h$ et $\sigma : h \mapsto (e_K, h)$.
Inversement, la proposition 5.2.1 peut s'exprimer de la façon suivante.

5.3.4. Corollaire. *Considérons une suite exacte courte de groupes (topologiques) :*

$$1 \longrightarrow K \xrightarrow{\iota} G \xrightarrow{\pi} \Gamma \longrightarrow 1.$$

Si π possède une section (continue), alors G est isomorphe à un produit semi-direct (topologique) $K \rtimes \Gamma$.

Autrement dit, s'il existe un morphisme de groupes (continu) $\sigma : \Gamma \to G$ tel que $\pi \circ \sigma = \mathrm{Id}_\Gamma$, alors G est isomorphe à un produit semi-direct (topologique) $K \rtimes_\varphi \Gamma$, où $\varphi_\gamma(k) = \sigma(\gamma) k \sigma(\gamma^{-1})$ pour $\gamma \in \Gamma$ et $k \in K$.

Démonstration. Soit $H = \sigma(\Gamma)$. Il suffit de vérifier les hypothèses de la proposition 5.2.1. On sait déjà que K est distingué[6] dans G. Si un élément h appartient à $K \cap H$, alors c'est l'image $\sigma(\gamma)$ d'un élément de Γ et son image par π est le neutre e_Γ de Γ (K est le noyau de π). D'où : $\gamma = \pi\sigma(\gamma) = \pi(h) = e_\Gamma$, puis : $h = \sigma(e_\Gamma)$ est le neutre de G. Enfin, soit g un élément de G. Alors, g et $\sigma\pi(g)$ ont la même image par π donc ces éléments diffèrent d'un élément du noyau : il existe k de K tel que $g = \sigma\pi(g)k$, si bien que g appartient à HK. On en tire : $G = HK$ puis, par passage à l'inverse : $G = KH$. □

5.3.5. Exemple. Soit $n \in \mathbb{N}^*$. Dans l'action naturelle de $\mathrm{GL}_{n+1}(\mathbb{R})$ sur l'espace $\mathbb{R}^{n+1}$, le stabilisateur G du premier vecteur de la base canonique est décrit par :

$$G = \left\{ \begin{pmatrix} 1 & x \\ 0 & g \end{pmatrix}, x \in \mathscr{M}_{1,n}(\mathbb{R}),\ g \in \mathrm{GL}_n(\mathbb{R}) \right\}.$$

On calcule, pour $x, x' \in \mathscr{M}_{1,n}(\mathbb{R})$ et $g, g' \in \mathrm{GL}_n(\mathbb{R})$:

$$\begin{pmatrix} 1 & x \\ 0 & g \end{pmatrix} \begin{pmatrix} 1 & x' \\ 0 & g' \end{pmatrix} = \begin{pmatrix} 1 & xg' + x' \\ 0 & gg' \end{pmatrix}.$$

On en déduit un morphisme de groupes :

$$\pi : G \to \mathrm{GL}_n(\mathbb{R}), \quad \begin{pmatrix} 1 & x \\ 0 & g \end{pmatrix} \mapsto g,$$

dont le noyau est un sous-groupe distingué, isomorphe à $\mathbb{R}^n$. On peut écrire une suite exacte :

$$1 \to \begin{pmatrix} 1 & \mathscr{M}_{1,n}(\mathbb{R}) \\ 0 & \mathrm{I}_n \end{pmatrix} \xrightarrow{\mathrm{Id}} G = \begin{pmatrix} 1 & \mathscr{M}_{1,n}(\mathbb{R}) \\ 0 & \mathrm{GL}_n(\mathbb{R}) \end{pmatrix} \xrightarrow{\pi} \mathrm{GL}_n(\mathbb{R}) \to 1.$$

Le morphisme π possède un section continue :

$$\sigma : \mathrm{GL}_n(\mathbb{R}) \longrightarrow G, \quad g \longmapsto \begin{pmatrix} 1 & 0 \\ 0 & g \end{pmatrix}.$$

Ainsi, le groupe G est naturellement isomorphe à un produit semi-direct $\mathscr{M}_{1,n}(\mathbb{R}) \rtimes \mathrm{GL}_n(\mathbb{R})$. Pour comprendre l'action du groupe linéaire décrite sur $\mathscr{M}_{1,n}(\mathbb{R})$, il suffit de conjuguer dans G. Pour $g \in \mathrm{GL}_n(\mathbb{R})$ et $x \in \mathscr{M}_{1,n}(\mathbb{R})$,

[6] C'est un abus de langage, il s'agit en fait de son image par ι.

on a :

$$\begin{pmatrix} 1 & 0 \\ 0 & g \end{pmatrix} \begin{pmatrix} 1 & x \\ 0 & \mathrm{I}_n \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & g \end{pmatrix}^{-1} = \begin{pmatrix} 1 & x \\ 0 & g \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & g^{-1} \end{pmatrix} = \begin{pmatrix} 1 & xg^{-1} \\ 0 & \mathrm{I}_n \end{pmatrix}.$$

L'action est donc définie par : $g \cdot x = xg^{-1}$ $\left(g \in \mathrm{GL}_n(\mathbb{R}),\ x \in \mathscr{M}_{1,n}(\mathbb{R})\right)$.

5.3.6. Exercice. *Étudier de même le stabilisateur du dernier vecteur de la base canonique pour l'action du groupe* $\mathrm{GL}_{n+1}(\mathbb{R})$ *sur* $\mathscr{M}_{n+1,1}(\mathbb{R})$ *définie par :* $g \cdot L = Lg^{-1}$ *pour* $g \in \mathrm{GL}_{n+1}(\mathbb{R})$ *et* $L \in \mathscr{M}_{n+1,1}(\mathbb{R})$. *Comparer avec l'exemple* 5.4.

Étudier de même, pour l'action de $\mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$ *sur l'espace des matrices* $\mathscr{M}_{m,n}(\mathbb{K})$, *le stabilisateur de la matrice* $\mathrm{I}_{m,n,r}$ *rencontré en page 8.*

5.4. L'exemple du groupe affine

Le *groupe affine* $\mathrm{GA}_n(\mathbb{K})$ est le groupe des bijections affines de l'espace affine $\mathbb{A}^n(\mathbb{K})$ identifié à $\mathbb{K}^n$. On peut le voir comme le groupe des bijections de degré 1 (mais pas nécessairement homogènes, sinon ce serait le groupe linéaire) de $\mathbb{K}^n$ dans lui-même.

On vérifie que l'on peut décrire matriciellement le groupe affine comme le groupe des matrices $(n+1)\times(n+1)$ de la forme

$$B = \begin{pmatrix} A & Y \\ 0 & 1 \end{pmatrix},\ A \in \mathrm{GL}_n(\mathbb{K}),\ Y \in \mathscr{M}_{n,1}(\mathbb{K}).$$

On identifie B à un élément du groupe affine en lui assignant la bijection

$$g_B(x_1, \cdots, x_n) = A\,{}^t(x_1, \cdots, x_n) + Y.$$

On a ainsi bien défini un isomorphisme de groupes entre les matrices de cette forme et le groupe des bijections affines de $\mathbb{A}^n$. La matrice A est alors la matrice, dans la base canonique de $\mathbb{K}^n$, de ce que l'on appelle l'application linéaire $\overrightarrow{g}_B$ associée à g_B. La matrice Y n'est rien d'autre que l'image de l'origine. On peut résumer cette propriété dans la proposition suivante.

5.4.1. Proposition. *Il existe une injection du groupe affine* $\mathrm{GA}_n(\mathbb{K})$ *vers le groupe linéaire* $\mathrm{GL}_{n+1}(\mathbb{K})$ *qui envoie l'automorphisme affine* g *de l'espace affine* $\mathbb{A}^n$ *(identifié à* $\mathbb{K}^n$*) sur la matrice*

$$\begin{pmatrix} A & Y \\ 0 & 1 \end{pmatrix},$$

où A *est la matrice de l'endomorphisme* $\overrightarrow{g}$ *dans la base canonique de* $\mathbb{K}^n$ *et où* Y *est le vecteur colonne image de l'origine par* g.

5.4.2. Remarque. Peut-être le lecteur scrupuleux voudra-t-il comprendre les raisons géométriques de cette représentation[7] du groupe affine $\mathrm{GA}_n(\mathbb{K})$. Comme il est souvent pratique de le faire —on aura tout le loisir de l'observer au cours des chapitres de géométrie projective— on peut identifier l'espace affine $\mathbb{A}^n(\mathbb{K})$ avec l'ensemble des droites vectorielles de $\mathbb{K}^{n+1}$ non incluses dans l'hyperplan $x_{n+1} = 0$ de $\mathbb{K}^{n+1}$. Une droite vectorielle sera ainsi identifiée à son point d'intersection avec l'hyperplan affine d'équation $x_{n+1} = 1$. L'image de l'application associée à l'action de $\mathrm{GL}_{n+1}(\mathbb{K})$ sur l'ensemble des droites de $\mathbb{K}^{n+1}$, c'est-à-dire sur l'espace projectif $\mathbb{P}^n(\mathbb{K})$, s'identifie, comme on l'a vu dans la proposition 4.3.6, au groupe projectif $\mathrm{PGL}_{n+1}(\mathbb{K}) = \mathrm{GL}_{n+1}(\mathbb{K})/\mathbb{K}^*$. Dans ce contexte, on voit le groupe affine comme le sous-groupe de $\mathrm{PGL}_{n+1}(\mathbb{K})$ qui laisse fixe l'hyperplan $x_{n+1} = 0$, que l'on peut considérer comme un *hyperplan à l'infini*. On peut alors relever ce sous-groupe en un sous-groupe de $\mathrm{GL}_{n+1}(\mathbb{K})$ qui laisse fixe la dernière coordonnée. Ce qui donne la représentation voulue.

On peut reformuler ceci en termes de suites exactes.

5.4.3. Exercice. *Montrer qu'on a la suite exacte :*

$$1 \longrightarrow \tau_n(\mathbb{K}) \longrightarrow \mathrm{GA}_n(\mathbb{K}) \underset{\text{section}}{\overset{\longleftarrow}{\longrightarrow}} \mathrm{GL}_n(\mathbb{K}) \longrightarrow 1.$$

Ici, $\tau_n(\mathbb{K})$ est le groupe des translations du $\mathbb{K}$-espace affine $\mathbb{A}^n$, isomorphe au groupe $\mathbb{K}^n$, et la surjection est l'application $g \mapsto \overrightarrow{g}$. Tout point O de l'espace affine définit une section s_O par : $s_O(\overrightarrow{g})(M) = O + \overrightarrow{g}(\overrightarrow{OM})$.

Par le corollaire 5.3.4, on obtient une description du groupe affine en termes de produit semi-direct :

$$\mathrm{GA}_n(\mathbb{K}) \simeq \tau_n(\mathbb{K}) \rtimes_{\varphi_O} \mathrm{GL}_n(\mathbb{K}),$$

où φ_O est donné par l'action par conjugaison de $s_O\big(\mathrm{GL}_n(\mathbb{K})\big)$ sur $\tau_n(\mathbb{K})$. Matriciellement, une fois choisie une origine O de $\mathbb{A}^n$, la section s_O se décrit par :

$$s_O : A \mapsto \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}.$$

Dans les notations du corollaire 5.3.4, H et K sont les sous-groupes formés des matrices de forme respectives

$$\begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix} \ (A \in \mathrm{GL}_n(\mathbb{K})) \quad \text{et} \quad \begin{pmatrix} \mathrm{I}_n & Y \\ 0 & 1 \end{pmatrix} \ (Y \in \mathcal{M}_{n,1}(\mathbb{K}) = \mathbb{K}^n).$$

[7] Il s'agit bien d'une représentation linéaire au sens propre du terme.

L'action par conjugaison se calcule sans mal :

$$\begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathrm{I}_n & Y \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \mathrm{I}_n & AY \\ 0 & 1 \end{pmatrix}.$$

Et donc, l'action φ_O de H sur K n'est rien d'autre que l'action naturelle du groupe linéaire $\mathrm{GL}_n(\mathbb{K})$ sur $\mathbb{K}^n$. *What else?*

Références

On trouvera des informations complémentaires dans : [25], chapitres 1 et 2 ; pour les produits semi-directs et les suites exactes : [23] et [29], chapitre 1.

A. Annexe. Normes sur les espaces de matrices

A.1. Topologie naturelle d'un espace normé et critère de continuité

Soit E un espace vectoriel réel ou complexe de dimension finie d. On rappelle ce qui suit.

- Toutes les normes sur E sont équivalentes ; elles induisent donc toutes la même topologie sur E, que l'on appelle *topologie naturelle*, et dont les ouverts sont les parties contenant une boule ouverte centrée en chacun de leurs points.
- Une fois choisie une base quelconque $(e_1, \ldots, e_d)$ de E et sa base duale $(x_1, \ldots, x_d)$, pour tout espace topologique Y, une application $f : Y \to E$ est continue si et seulement si $x_i \circ f$ est continue pour tout entier i compris entre 1 et d.
- Avec des arguments standard, on en déduit que la somme ou le produit de deux applications continues à valeurs dans une algèbre normée de dimension finie (une algèbre de matrices par exemple) est continue. L'inverse d'une fonction continue à valeurs non nulles reste continue.
- Sur un espace de dimension finie E, toute forme linéaire est continue. Donc, en particulier, par le point qui précède, il en est de même de toute fonction polynomiale sur E, *i.e.* de toute fonction polynomiale en les coordonnées d'un vecteur (cette caractérisation ne dépend pas de la base choisie). De même, une fraction rationnelle est continue sur toute partie où elle est définie.

A.1.1. Remarque. En particulier, quelle que soit la taille des matrices, la multiplication matricielle $\mathscr{M}_{m,n}(\mathbb{K}) \times \mathscr{M}_{n,p}(\mathbb{K}) \to \mathscr{M}_{m,p}(\mathbb{K})$ est continue. De même, le passage à l'inverse $\mathrm{GL}_n(\mathbb{K}) \to \mathrm{GL}_n(\mathbb{K})$, $g \mapsto g^{-1}$ est continu, par la formule de la comatrice.

À titre d'illustration, voici une application très utile de ce critère pour la topologie d'espace de matrices.

A.1.2. Proposition. *Soit n un entier naturel, $\mathscr{M}_n(\mathbb{K})$ l'algèbre des matrices $n\times n$ à coefficients dans $\mathbb{K}$ et $\mathbb{K}_n[\lambda]$ l'espace des polynômes en une indéterminée λ de degré inférieur ou égal à n. L'application « polynôme caractéristique » $\chi : \mathscr{M}_n(\mathbb{K}) \to \mathbb{K}_n[\lambda]$, $A \mapsto \chi_A$, est continue.*

Démonstration. En effet, chacun des coefficients du polynôme caractéristique est une fonction polynomiale des coefficients de la matrice. Or, les coefficients d'un polynôme sont ses coordonnées dans la base canonique de $\mathbb{K}_n[\lambda]$. □

A.1.3. Remarque. Voir le second tome, exercice « Mineurs et polynôme caractéristique » dans § I-2.2 pour la description de ces polynômes.

A.1.4. Exercice. *Montrer que l'ensemble des matrices de $\mathscr{M}_n(\mathbb{C})$ à valeurs propres distinctes est un ouvert de $\mathscr{M}_n(\mathbb{C})$.*
On pourra considérer l'application continue qui à la matrice A associe le discriminant de χ_A, c'est-à-dire le résultant de χ_A et de χ'_A.

A.2. Normes subordonnées

Soit n un entier naturel et $\mathscr{M}_n(\mathbb{K})$ l'espace des matrices carrées $n \times n$ à coefficients dans $\mathbb{K}$. Parmi toutes les normes figurent les *normes subordonnées* à une norme de $\mathbb{K}^n$. Étant donné une norme $\|\cdot\|$ sur $\mathbb{K}^n$, on lui associe une norme $|||\cdot|||$ sur $\mathscr{M}_n(\mathbb{K})$ par :

$$\forall A \in \mathscr{M}_n(\mathbb{K}), \quad |||A||| = \sup_{\|x\|=1} \|Ax\| .$$

L'avantage d'une norme subordonnée, c'est qu'elle est multiplicative :

$$\forall A, A' \in \mathscr{M}_n(\mathbb{K}), \quad |||AA'||| \leqslant |||A|||\times|||A'|||.$$

Et puisque toutes les normes sont équivalentes, on peut se permettre d'en prendre une qui possède de (plus) belles propriétés ! En voici quelques exemples :

norme vectorielle	norme matricielle
$\|x\|_1 = \sum_{1\leqslant i\leqslant n} \lvert x_i\rvert$	$\|\|\|A\|\|\|_1 = \max_{1\leqslant j\leqslant n} \sum_{1\leqslant i\leqslant n} \lvert a_{ij}\rvert$
$\|x\|_2 = \sqrt{\sum_{1\leqslant i\leqslant n} \lvert x_i\rvert^2}$	$\|\|\|A\|\|\|_2 = \sqrt{\rho(A^*A)}$
$\|x\|_\infty = \max_{1\leqslant i\leqslant n} \lvert x_i\rvert$	$\|\|\|A\|\|\|_\infty = \max\limits_{1\leqslant i\leqslant n} \sum_{1\leqslant j\leqslant n} \lvert a_{ij}\rvert$

Dans le tableau, ρ désigne le rayon spectral : pour B matrice carrée, $\rho(B)$ est le module maximal d'une valeur propre de B. On prouve facilement la première et la troisième ligne grâce à une double inégalité. En ce qui concerne la deuxième ligne, on prouvera le résultat qui y figure plus tard

en exercice comme application naturelle de la décomposition polaire, voir le corollaire VI-1.6.

A.3. Aparté : axiomes de séparation

Un espace X est dit *séparé* si, pour tous x et x' distincts dans X, il existe un voisinage V de x et un voisinage V' de x' qui ne se coupent pas.
Il existe d'autres axiomes de séparation : un espace X est dit T_1, ou séparé au sens de Fréchet, si pour tous x et x' distincts dans X, il existe un voisinage V de x qui ne contient pas x'. Il est équivalent de demander que les singletons soient fermés (vérifier).
Pour les groupes topologiques, les deux notions coïncident.

A.3.1. Lemme. *Soit G un groupe topologique, soit e son neutre. Il est équivalent de dire :*

(i) *G est séparé ;*

(ii) *pour tout g de G, le singleton $\{g\}$ est fermé dans G ;*

(iii) *le singleton $\{e\}$ est fermé dans G.*

Démonstration. Si G est séparé, il est T_1 et donc les singletons sont fermés.
Réciproquement, supposons que les singletons soient fermés. Soit g un point différent de e. L'ouvert $G \setminus \{g\}$ contient e donc son image réciproque par le produit $\mu : G \times G \to G$ est un ouvert contenant (e, e). Par définition de la topologie produit, il existe deux ouverts V_1 et V_2 de G, qui contiennent e, tels qu'on ait : $V_1 \times V_2 \subset \mu^{-1}(G \setminus \{e\})$. Soit $V = V_1 \cap V_2$, on a : $VV = \mu(V, V) \subset G \setminus \{e\}$. Quitte à remplacer V par $V \cap V^{-1}$, on peut supposer de plus que $V = V^{-1}$: c'est (toujours) un voisinage ouvert de e. Mais alors, gV et V sont deux ouverts disjoints contenant respectivement g et g. On en déduit par le principe de translation que G est séparé.
On vient de démontrer l'équivalence de (i) et (ii). L'équivalence entre (ii) et (iii) résulte du principe de translation : pour tout élément g de G, la translation $h \mapsto gh$ est un homéomorphisme et elle envoie le fermé $\{e\}$ sur le singleton $\{g\}$, donc ce dernier est fermé. □

B. Annexe. Connexité

Définitions de la connexité

B.1. Lemme. *Soit X un espace topologique. Les conditions suivantes sont équivalentes :*

(i) *les seules parties ouvertes et fermées de X sont X et $\emptyset$;*

(ii) *on ne peut pas écrire X comme la réunion de deux fermés disjoints non vides ;*

(iii) *toute fonction continue de X vers $\{0, 1\}$ est constante ;*

(iv) *toute fonction continue de X vers un ensemble discret est constante.*

B.2. Remarque. L'assertion (iii) est très commode en pratique pour prouver la connexité d'un espace. (Il va sans dire que la topologie de $\{0,1\}$ est la topologie discrète.)

B.3. Définition. Soit X un espace topologique. On dit que X est *connexe* si les conditions équivalentes du lemme sont satisfaites.

B.4. Définition. Soit X un espace topologique. On dit que X est *connexe par arcs* si, pour tout couple de points (x,y) de X, il existe une application continue $\gamma : [0,1] \to X$ telle que $\gamma(0) = x$ et $\gamma(1) = y$.

Critères de connexité

B.5. Proposition. *Soit X et Y deux espaces topologiques.*

(i) *Soit $f : X \to Y$ une fonction continue surjective. Si X est connexe (resp. connexe par arcs), alors Y l'est également.*

(ii) *Si X et Y sont connexes, alors le produit cartésien $X \times Y$, muni de la topologie produit, est connexe.*

(iii) *Si X est connexe par arcs, alors il est connexe.*

B.6. Remarque. En général, la réciproque de l'assertion (iii) est fausse : il existe des espaces connexes qui ne sont pas connexes par arcs, par exemple le « peigne » $\big((\mathbb{Q} \cap [0,1[) \times [0,1]\big) \cup \{(1,1)\}$.

Cependant, les ouverts d'un espace normé, plus généralement, les variétés différentielles et, en particulier, les groupes de Lie, que l'on rencontrera au chapitre IX, constituent une classe d'espaces pour lesquels la connexité et la connexité par arcs sont équivalentes.

B.7. Proposition. *Soit Z un espace topologique.*

(i) *Soit X et Y deux parties de Z telles que X soit connexe et que l'on ait : $X \subset Y \subset \overline{X}$. Alors, Y est connexe.*

(ii) *Soit $(X_1, \ldots, X_n)$ une famille finie de parties connexes de Z. Si l'intersection de X_i et X_{i+1} n'est pas vide ($1 \leqslant i < n$), alors la réunion des X_i est connexe.*

(iii) *Soit $(X_i)_{i \in I}$ une famille quelconque de parties connexes de Z dont l'intersection n'est pas vide. Alors, la réunion des X_i est connexe.*

Composantes connexes

B.8. Définition. Soit X un espace topologique. On définit une relation $\sim$ sur X de la façon suivante : pour x et y éléments de X, on dit que $x \sim y$ s'il existe une partie connexe qui contient x et y. Cette relation est une relation d'équivalence (la transitivité vient de la proposition B.7 (ii)). On appelle *composantes connexes* de X les classes d'équivalence de la relation $\sim$.

B.9. Lemme. *Les composantes connexes d'un espace topologique sont les parties connexes maximales et elles sont fermées.*

Démonstration. D'après la proposition B.7 (iii), la réunion de toutes les parties connexes contenant un point est connexe et c'est manifestement une partie connexe maximale. D'après la proposition B.7 (i), l'adhérence d'une partie connexe est connexe. Par maximalité, une composante connexe est donc fermée. □

B.10. Remarque. Si l'espace topologique X est réunion disjointe d'ouverts connexes O_i, alors chaque O_i est aussi fermé, comme complémentaire d'ouvert. Par conséquence, toute partie contenant strictement O_i est non connexe (union de deux ouverts fermés disjoints). On peut faire de même en remplaçant O_i par des fermés F_i à condition d'en prendre un nombre fini. Attention, cela devient faux si l'on ne précise pas « fini » ; prendre tous les points fermés connexes de la droite réelle, ce ne sont pas des composantes connexes...

C. Annexe. Compacité locale

C.1. Définition. Un espace topologique est dit *localement compact* s'il est séparé et si tout point possède un voisinage ouvert dont l'adhérence est compacte.

On trouve aussi dans la littérature la définition suivante.

C.2. Définition (variante). Un espace topologique est dit *localement compact* s'il est séparé et si tout point possède une base de voisinages ouverts à adhérence compacte.

Le fait que les deux définitions sont équivalentes repose sur la séparation.

C.3. Lemme. *Les deux définitions précédentes sont équivalentes.*

Démonstration. Il suffit de démontrer que la première définition implique la seconde. Soit donc X un espace localement compact au sens de la première définition et x un point de X. Soit V un voisinage ouvert de x. On doit prouver l'existence d'un voisinage compact de x contenu dans V. Soit U un voisinage ouvert de x dont l'adhérence est compacte ; quitte à remplacer U

par $U \cap V$, on peut supposer que U est inclus dans V. Supposons avoir construit un voisinage ouvert W de x inclus dans U et dont l'adhérence $\overline{W}$ est aussi incluse dans U. Alors, $\overline{W}$ serait compact et inclus dans V, ce qui entraînerait le lemme.

Prouvons donc l'existence d'un tel W. Soit ∂U la frontière de U ; c'est un compact. Pour tout y de ∂U, il existe des voisinages ouverts U_y de y et W_y de x qui les séparent. Du recouvrement du compact ∂U par les ouverts U_y, on extrait un sous-recouvrement fini $(U_{y_1}, \ldots, U_{y_N})$. Soit alors W l'intersection finie des W_{y_i} $(1 \leqslant i \leqslant N)$: c'est un voisinage ouvert de x, inclus dans U, et qui, par construction, n'intersecte pas l'ouvert $\cup_i U_{y_i}$ recouvrant ∂U. Par suite, tout élément de ∂U a un voisinage ouvert ne rencontrant pas W (il suffit de prendre un U_{y_i} auquel il appartient) et n'appartient donc pas à ∂W.

On a donc bien : $\overline{W} \subset U$, ce qui était notre objectif. □

Le corps des réels, celui des complexes et, plus généralement, les espaces de dimension finie sur ces corps sont localement compacts pour la topologie normique. Cela provient du fait qu'un compact est un fermé borné.

Si X est un espace localement compact, tout ouvert V de X l'est également. En effet, tout ouvert de V est aussi un ouvert de X ; on se ramène alors à la définition de la compacité en termes de recouvrements ouverts.

Si X est un espace localement compact, tout fermé F de X l'est également. En effet, l'intersection d'un fermé et d'un compact est compacte (c'est un fermé dans un compact).

On déduit de ces constats la propriété suivante.

C.4. Proposition. *Soit X un espace topologique localement compact. Si $X_n \subset X_{n-1} \subset \cdots \subset X_0 = X$ est une suite finie de parties telles que X_i est ouvert ou fermé dans X_{i-1} pour i compris entre 1 et n, alors X_n est localement compact.*

La preuve du contre-exemple ci-dessous n'est pas très efficace, mais elle est instructive.

C.5. Lemme. *L'espace métrique $\mathbb{Q}$ n'est pas localement compact.*

Démonstration. Supposons, au contraire, que $\mathbb{Q}$ soit localement compact. Alors, 0 possède un voisinage ouvert U dont l'adhérence est compacte. Tout ouvert de U a aussi une adhérence compacte donc on peut choisir pour U un intervalle de la forme $]-t, t[\cap \mathbb{Q}$, pour t irrationnel strictement positif. (Ces parties forment bien une base d'ouverts de 0 dans $\mathbb{Q}$.) Comme t est irrationnel, la partie $]-t, t[\cap \mathbb{Q} = [-t, t] \cap \mathbb{Q}$ est ouverte et fermée dans $\mathbb{Q}$. Par suite, U est compact. Choisissons à présent une suite $(x_n)_{n \geqslant 0}$ de rationnels appartenant à U qui converge vers t. On obtient une suite

à valeurs dans un compact qui ne possède pas de sous-suite convergente dans le compact en question (toutes les sous-suites tendent vers t). Ceci contredit le théorème de Bolzano-Weierstrass, d'où le lemme. □

Et voici maintenant de quoi briller en société.

C.6. Définition. Un espace topologique est appelé *espace de Baire* si toute intersection dénombrable d'ouverts denses est dense. En prenant le complémentaire, cela donne la définition équivalente que si la réunion dénombrable de fermés F_i est d'intérieur non vide, alors un des F_i est d'intérieur non vide.

Notre besoin des espaces localement compacts tient largement au théorème suivant.

C.7. Théorème. *Tout espace topologique localement compact est un espace de Baire.*

Démonstration. Soit X un espace localement compact et $(U_i)_{i\in\mathbb{N}}$ une suite d'ouverts partout denses. Pour prouver que leur intersection est partout dense, il suffit de prouver que si V est un ouvert quelconque, il existe un point appartenant à V et à tous les U_i. Nous allons construire une suite d'ensembles compacts $(B_j)_{j\in\mathbb{N}}$ satisfaisant à : $B_1 \subset U_1 \cap V$ et, pour $j \geqslant 0$: $B_{j+1} \subset U_{j+1} \cap \overset{\circ}{B}_j$. Il suffira alors de prouver que l'intersection des B_j n'est pas vide pour terminer la preuve.

Choisissons un point x de $U_1 \cap V$. Comme X est localement compact, x possède un voisinage B_1 compact et contenu dans $U_1 \cap V$. On construit de même B_{j+1} à partir de $U_{j+1} \cap \overset{\circ}{B}_j$. Comme une suite décroissante de compacts non vide n'est pas vide (c'est une conséquence de la propriété de Borel-Lebesgue), l'intersection des B_j n'est pas vide. Un point de cette intersection appartient à V et à tous les U_j, d'où le théorème. □

C.8. Remarque. Voici une preuve express du fait que $\mathbb{Q}$ n'est pas localement compact : $\mathbb{Q}$ n'est pas un espace de Baire, car l'intersection des ouverts denses $\mathbb{Q} \setminus \{r_i\}$ $(i \in \mathbb{N})$ est vide, où $(r_i)_{i\in\mathbb{N}}$ est une énumération de tous les rationnels.

Notons que l'ensemble des irrationnels n'est pas localement compact alors que c'est un espace de Baire. Notons pour finir que la compacité locale est essentielle pour définir la compactification d'Alexandroff, c'est-à-dire la possibilité de prolonger l'espace X en un espace $\overline{X}$ qui est compact.

D. Annexe. Topologie sur un espace quotient

Soit G un groupe topologique et X un espace topologique muni d'une action continue de G. Pour g dans G et x dans X, on note $g{\cdot}x$ le résultat de l'action de g sur x. On suppose que G et X sont séparés.

Le cours regorge d'exemples où l'action de G sur X est transitive : dans les bons cas, on sait que X est homéomorphe à un quotient G/H, où H est le stabilisateur d'un point de X. Mais il arrive que l'ensemble des orbites ait un intérêt géométrique intrinsèque, alors que l'action n'est pas transitive. Il est alors utile de mettre une topologie sur l'ensemble X/G des orbites des éléments de X sous l'action de G.

D.1. Topologie quotient sur X/G : définition et propriétés immédiates

D.1.1. Lemme. *Soit G un groupe topologique agissant continûment sur un espace topologique X. Les parties de X/G dont les images réciproques par la projection naturelle $\pi : X \to X/G$ sont ouvertes dans X forment une topologie. C'est la topologie la plus fine qui rend la projection π continue.*

Démonstration. Si $(\overline{S_i})_{i\in I}$ est une famille de telles parties, la réunion des $\pi^{-1}(\overline{S_i})$ est ouverte en tant que réunion d'ouverts, si bien que la réunion des $\overline{S_i}$ est bien dans la topologie ; si la famille est finie, leur intersection est ouverte comme intersection finie d'ouverts donc l'intersection des $\overline{S_i}$ est aussi dans la topologie ; enfin, $\emptyset = \pi^{-1}(\emptyset)$ et $X = \pi^{-1}(X/G)$.

Pour la deuxième assertion, constatons que dans n'importe quelle topologie pour laquelle π est continue, si $\overline{S}$ est une partie ouverte de X/G, $\pi^{-1}(\overline{S})$ est ouvert dans X. Ici, on prend toutes les parties satisfaisant à cette contrainte. □

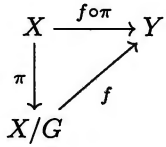
D.1.2. Définition (Topologie quotient sur X/G). Soit G un groupe topologique agissant continûment sur un espace topologique X. La *topologie quotient* sur X/G est la topologie dont les ouverts sont les parties dont l'image réciproque par la projection canonique $\pi : X \to X/G$ sont ouvertes.

Par construction, π est ouverte : si U est un ouvert de X, alors $\pi(U)$ est ouvert dans X/G. En effet, la saturation de U, $\pi^{-1}\big(\pi(U)\big)$, est la réunion des $gU = \tau_{g^{-1}}^{-1}(U)$ lorsque g parcourt G, où $\tau_{g^{-1}} : X \to X$, $x \mapsto g^{-1}\cdot x$ est l'action de g^{-1} sur X, une application continue.

Voici donc la proposition attendue.

D.1.3. Proposition. *Soit G un groupe topologique agissant continûment sur un espace topologique X ; soit Y un espace topologique et $f : X/G \to Y$ une application. Alors, f est continue si et seulement si $f \circ \pi$ est continue.*

Démonstration. En effet, considérons le diagramme :



Pour W un ouvert de Y, $(f \circ \pi)^{-1}(W) = \pi^{-1}\big(f^{-1}(W)\big)$ est ouvert si et seulement si $f^{-1}(W)$ est ouvert (on voit là l'intérêt d'avoir pris la topologie quotient aussi fine que possible).

Enfin, si Y est un espace topologique et $F : X \to Y$ est une application continue et constante sur chaque orbite, l'application $f : X/G \to Y$ induite par passage au quotient est continue. En effet, on a par définition de f la relation $F = f \circ \pi$; pour W ouvert de Y, $f^{-1}(W)$ est ouvert dans X/G, car $\pi^{-1}\big(f^{-1}(W)\big) = F^{-1}(W)$ est ouvert dans X. □

D.2. Exemple : espaces projectifs

Définition. L'ensemble des droites vectorielles d'un espace vectoriel E de dimension n sur un corps $\mathbb{K}$ est appelé espace *projectif* (de dimension $n-1$) et noté $\mathbb{P}(E)$ ou $P^{n-1}(\mathbb{K})$ si E est assimilé à l'espace $\mathbb{K}^n$.

On a vu au § 4.3 que cet ensemble s'identifie au quotient de $\mathrm{GL}_n(\mathbb{K})$ par un sous-groupe fermé, ce qui permet de lui donner une topologie séparée. Mais cette réalisation n'est pas naturelle : d'une part, elle dépend du choix d'un vecteur non nul ; mais, plus subjectivement, on pense plus volontiers à une droite comme à une collection de vecteurs colinéaires que comme une classe de matrices inversibles modulo un énorme sous-groupe P...

Aussi est-il plus naturel de définir l'espace projectif comme l'ensemble-quotient de $X = E \setminus \{0\}$ par l'action du groupe G des scalaires non nuls.

Ce qui, en passant, nous permet de définir des sous-espaces projectifs : on appellera droite (projective) de $P^{n-1}(\mathbb{K})$ l'image d'un plan de $\mathbb{K}^n$ dans le quotient, et hyperplan de $P^{n-1}(\mathbb{K})$ l'image d'un hyperplan de E.

Si le corps est celui des réels ou celui des complexes, il est naturel de définir une topologie sur ce quotient, quitte à la comparer le plus tôt possible à celle de $\mathrm{GL}_n(\mathbb{K})/P$ – voir l'exercice F.23.

D.3. Lien avec les espaces homogènes G/H

Pour garder des notations localement cohérentes, l'espace homogène s'appellera ici $\tilde{G}/G$ et pas G/H. Soit donc $\tilde{G}$ un groupe topologique et G un sous-groupe fermé de $\tilde{G}$, on fait agir G sur $X = \tilde{G}$ par translations à droite :

$$\begin{aligned} \phi : G \times \tilde{G} &\longrightarrow \tilde{G} \\ (g, \gamma) &\longmapsto g \cdot \gamma = \gamma g^{-1}. \end{aligned}$$

Alors, l'orbite d'un élément γ de $\tilde{G}$ est exactement sa classe γG, si bien que la notation $\tilde{G}/G$, qui désigne le quotient d'un groupe par un sous-groupe ou l'ensemble des orbites d'un groupe agissant sur un ensemble, n'est pas ambiguë.

Mieux : la topologie quotient que l'on vient de mettre sur l'ensemble des orbites X/G avec $X = \tilde{G}$ est bien la topologie du quotient $\tilde{G}/G$ du groupe $\tilde{G}$ par son sous-groupe G. On généralise donc bien ce qui a été fait sur les groupes quotients.

On peut alors faire agir $\tilde{G}$ sur $\tilde{G}/G$ par translation à gauche : si γG désigne la classe de $\gamma \in \tilde{G}$ et si $\gamma' \in \tilde{G}$, on définit l'action de γ' sur γG comme la classe $\gamma'\gamma G$. On vérifie que cela définit une action continue (ah ?) et transitive. D'évidence, le stabilisateur de la classe du neutre $eG \in \tilde{G}/G$ est le groupe G. Le théorème d'homéomorphisme est passionnant dans ce cas : $\tilde{G}/G$ est homéomorphe à $\tilde{G}/G$... Si l'on change de point, ce n'est guère mieux : le stabilisateur de γG est $G^\gamma = \gamma G \gamma^{-1}$, on obtient le fait que $\tilde{G}/G^\gamma$ est homéomorphe à $\tilde{G}/G$: la conjugaison par γ induit un tel homéomorphisme...

D.4. Cas des actions propres

On dit que l'action de G sur X est *propre* si, pour tout compact K de X, l'ensemble des éléments g de G tels que l'intersection de gK et K soit non vide est fini :

$$\forall K \subset X,\ K \text{ compact}, \quad \#\{g \in G,\ gK \cap K \neq \emptyset\} < +\infty.$$

Par exemple, l'action de $\mathbb{R}^*$ sur $\mathbb{R}^n \setminus \{0\}$ n'est pas propre (pourquoi ?). L'action de $\mathbb{Z}^2$ sur $\mathbb{R}^2$ par translation est, elle, propre. La séparation est, en quelque sorte, le minimum que l'on puisse espérer pour une topologie quotient ; voici un critère.

D.4.1. Proposition. *Si l'action de G sur X est continue et propre, et si G et X sont séparés et localement compacts, alors X/G est séparé.*

Démonstration. Fixons deux éléments distincts $\overline{x}$ et $\overline{y}$ distincts dans X/G, x et y des représentants. On cherche des ouverts disjoints U_0 et V_0 contenant

respectivement x et y, tels que V_0 n'intersecte aucun des translatés gU_0, $g \in G$: les images $\pi(U_0)$ et $\pi(V_0)$ seront alors des ouverts séparant $\overline{x}$ et $\overline{y}$. Vu les hypothèses faites sur X, on peut trouver U et V des ouverts disjoints de X contenant respectivement x et y, d'adhérence compacte. L'ensemble A des éléments g de G tels que $gU \cap V$ n'est pas vide est fini. En effet, s'il contient un élément g_0 et si g est un élément de G, on a :

$$gU \cap V \subset \left(gg_0^{-1} \cdot (g_0 U)\right) \cap V \subset \left(gg_0^{-1} \cdot K\right) \cap K,$$

où $K = g_0\overline{U} \cup \overline{V}$ est compact ; l'intersection ci-dessus est donc vide sauf pour un nombre fini de g.

Si A est vide, les ouverts U et V conviennent. Sinon, numérotons les éléments de A : $A = \{g_1, \ldots, g_n\}$, $n \in \mathbb{N}^*$. Soit i un entier compris entre 1 et n. Puisque $\overline{x}$ et $\overline{y}$ sont distincts, $g_i x$ et y le sont aussi, si bien que l'on peut trouver des ouverts disjoints $U_i' \subset g_i U$ et $V_i' \subset V$ contenant respectivement $g_i x$ et y. Notons alors :

$$U_0 = \bigcap_{i=1}^{n} g_i^{-1} U_i' \quad \text{et} \quad V_0 = \bigcap_{i=1}^{n} V_i.$$

Ces ouverts sont disjoints et contiennent respectivement x et y, et l'on a bien pour $g \in G$: $(gU_0) \cap V_0 = \emptyset$: en effet, c'est vrai si l'on a $g \notin A$ puisque $V_0 C V$, $U_0 C U$ et $V \cap gU = \emptyset$ (puisque $g \notin A$), et si $g = g_i \in A$, c'est vrai car $(g_i U_0) \cap V_0 C U_i' \cap V_i' = \emptyset$. Cela prouve que X/G est séparé.□

D.4.2. Exemple (du demi-plan de Poincaré aux surfaces de Riemann). Faisons opérer classiquement $\mathrm{SL}_2(\mathbb{R})$ par homographies sur le demi-plan de Poincaré $\mathscr{H} = \{z \in \mathbb{C},\ \mathrm{Im} z > 0\}$ (voir l'exercice F.18) :

$$\forall A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R}),\ \forall z \in \mathscr{H}, \quad A \cdot z = \frac{az+b}{cz+d}\,.$$

Soit alors $G = \mathrm{SL}_2(\mathbb{Z})$, le groupe des matrices 2×2 à coefficients entiers de déterminant 1. Prouver que l'action est propre revient à remarquer que l'intersection d'une partie discrète et d'une partie compacte est finie et que les stabilisateurs dans G des points de $\mathscr{H}$ sont finis.

On peut montrer que le quotient $\mathscr{H}/G$ paramètre les classes de *réseaux* dans $\mathbb{C}$ (sous-groupes isomorphes à $\mathbb{Z}^2$), voir l'exercice F.24 ou [34], chapitre VII, § 2.2. On obtient donc une topologie sur l'ensemble des réseaux ou même, en raffinant, une structure analytique : on repère un réseau par une coordonnée complexe, ce qui permet d'étudier « la géométrie de l'espace des réseaux ».

On peut énoncer une vague mais spectaculaire classification des surfaces (sans donner la définition d'une surface...) : toute surface compacte sans bord est homéomorphe à la sphère euclidienne, au tore $\mathbb{R}^2/\mathbb{Z}^2$ ou à un quotient $\mathscr{H}/\Gamma$, où Γ est un sous-groupe discret de $\mathrm{SL}_2(\mathbb{R})$.

E. Développement. Actions classiques et leurs invariants

Les tableaux qui suivent présentent, sans intention d'exhaustivité, des groupes classiques, et des actions classiques pour lesquelles on dispose d'invariants totaux. Le premier tableau relève de l'algèbre linéaire, le second plutôt de géométrie (même si la différence entre les deux n'est pas si claire). Ceci peut apporter un peu de recul sur les définitions introduites dans le programme de mathématiques. Enfin, le dernier tableau donne une liste d'actions transitives liées au théorème de la base incomplète.

Groupe G	Ensemble X	Action $G \times X \longrightarrow X$	Espace quotient[8] X/G	Invariant
$\mathrm{GL}_n(\mathbb{K})$	sous-espaces vectoriels de $\mathbb{K}^n$	$g \cdot F = g(F)$	$\{0, \dots, n\}$	dimension
$\mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$	matrices rectangulaires : $\mathscr{M}_{m,n}(\mathbb{K})$	$(P,Q) \cdot A = PAQ^{-1}$	$\{0, \dots, \min(m,n)\}$	rang, p. 3
$\mathrm{GL}_m(\mathbb{K})$	matrices rectangulaires : $\mathscr{M}_{m,n}(\mathbb{K})$	$P \cdot A = PA$	$\bigcup_{1 \leqslant i_1 < i_2 \cdots < i_m \leqslant n} \mathbb{K}^{\sum (i_k - 1)}$	matrices échelonnées réduites
$\mathrm{GL}_n(\mathbb{K})$	matrices diagonalisables : $\mathscr{D}_n(\mathbb{K})$	$P \cdot A = PAP^{-1}$	$\mathbb{K}^n / \mathfrak{S}_n$	valeurs propres, p. 84
$\mathrm{GL}_n(\mathbb{K})$	matrices nilpotentes : $\mathscr{N}_n(\mathbb{K})$	$P \cdot A = PAP^{-1}$	partitions de n	diagrammes de Young, p. 87
$\mathrm{GL}_n(\mathbb{K})$	matrices carrées : $\mathscr{M}_n(\mathbb{K})$	$P \cdot A = PAP^{-1}$	$P_1 \vert P_2 \vert \cdots \vert P_n$ polynômes unitaires	invariants de similitude, p. 83
$\mathrm{GL}_n(\mathbb{C})$	matrices symétriques : $\mathscr{S}_n(\mathbb{C})$	$P \cdot A = PA\,{}^tP$	$\{0, \dots, n\}$	rang, p. 151
$\mathrm{GL}_n(\mathbb{R})$	matrices symétriques : $\mathscr{S}_n(\mathbb{R})$	$P \cdot A = PA\,{}^tP$	$\{(p,q,r) \in \mathbb{N}^3 : p+q+r = n\}$	signature et rang, p. 151
$\mathrm{GL}_n(\mathbb{R})$	matrices symétriques inversibles : $\mathscr{S}_n(\mathbb{R}) \cap \mathrm{GL}_n(\mathbb{R})$	$P \cdot A = PA\,{}^tP$	$\{(p,q) \in \mathbb{N}^2 \ , \ p+q = n\}$	signature
$\mathrm{GL}_n(\mathbb{F}_q)$	matrices symétriques : $\mathscr{S}_n(\mathbb{F}_q)$	$P \cdot A = PA\,{}^tP$	$\{0\} \cup \{1, \dots, n\} \times \mathbb{Z}/2\mathbb{Z}$	rang et discriminant p. 151
$\mathrm{GL}_{2n}(\mathbb{K})$	matrices antisymétriques : $\mathscr{A}_{2n}(\mathbb{K})$	$P \cdot A = PA\,{}^tP$	$\{0, 2, \dots, 2n\}$	rang, p. 160
$\mathrm{GL}_n(\mathbb{Z})$	sous-réseaux de $\mathbb{Z}^n$	$g \cdot \mathscr{R} = g(\mathscr{R})$	$0 \leqslant r \leqslant n$ $d_1 \vert d_2 \vert \cdots \vert d_r \in \mathbb{N}$	rang et invariants de la base adaptée

[8] Ou ensemble de paramètres.

Groupe G	Ensemble X	Action $G \times X \longrightarrow X$	Espace quotient[9] X/G	Invariant
$\mathbb{K}^*$	vecteurs non nuls : $\mathbb{K}^{n+1} \setminus \{0\}$	$\lambda \cdot v = \lambda v$	$\mathbb{P}^n(\mathbb{K})$	droites de $\mathbb{K}^{n+1}$
$\mathrm{SO}_2(\mathbb{R})$	couples de droites du plan	action diagonale	$\mathbb{R}/\pi\mathbb{R}$	angles de droites
$\mathrm{SO}_2(\mathbb{R})$	couples de vecteurs de norme 1	$g \cdot (v, v') = (gv, gv')$	$\mathbb{R}/2\pi\mathbb{R}$	angles orientés de vecteurs
$\mathrm{O}_2(\mathbb{R})$	couples de droites du plan	$g \cdot (v, v') = (gv, gv')$	$\mathbb{R}/2\pi\mathbb{R} \mod x = -x$	angles non orientés de vecteurs
$\mathrm{PGL}_2(\mathbb{K})$	quadruplets de points de $\mathbb{P}^1(\mathbb{K})$ trois premiers distincts	action diagonale $A \cdot z = \dfrac{az+c}{bz+d}$	$\mathbb{P}^1(\mathbb{K})$	birapport cf p. 341 $\dfrac{z_4 - z_2}{z_4 - z_1} : \dfrac{z_3 - z_2}{z_3 - z_1}$
$\mathrm{GA}_2(\mathbb{R})$	coniques du plan	$g \cdot \mathscr{C} = g(\mathscr{C})$	classif. cf p. 169	genre des coniques
Is^+	coniques à centre	$g \cdot \mathscr{E} = g(\mathscr{E})$	$\{(a,b) \in \mathbb{N}^2 : a \leqslant b\}$	petit et grand axes, p. 164
$\mathbb{R}^* \times \mathrm{Is}^+$	coniques non dégénérées	$g \cdot \mathscr{E} = g(\mathscr{E})$	$]0, +\infty[$	excentricité

Ce tableau donne des exemples de formes normales que l'on rencontre tout au long du livre.

Groupe G	Ensemble X	Action $G \times X \longrightarrow X$	Forme normale
$\mathrm{GL}_n(\mathbb{K})$	sous-espaces vectoriels de $\mathbb{K}^n$	$g \cdot F = g(F)$	$F_0 = \{(x_i) \in \mathbb{K},\ x_{m+1} = \cdots = x_n = 0\}$
$\mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$	matrices rectangulaires : $\mathscr{M}_{m,n}(\mathbb{K})$	$(P,Q) \cdot A = PAQ^{-1}$	$\mathrm{I}_{m,n,r}$, p. 3
$\mathrm{GL}_m(\mathbb{K})$	matrices rectangulaires : $\mathscr{M}_{m,n}(\mathbb{K})$	$P \cdot A = PA$	$E \in \mathscr{M}_{m,n}(\mathbb{K})$, th. IV-2.3
$\mathrm{GL}_n(\mathbb{K})$	matrices diagonalisables : $\mathscr{D}_n(\mathbb{K})$	$P \cdot A = PAP^{-1}$	matrice diagonale (mod. $\mathfrak{S}_n$)
$\mathrm{GL}_n(\mathbb{K})$	matrices nilpotentes : $\mathscr{N}_n(\mathbb{K})$	$P \cdot A = PAP^{-1}$	J_ν, $\nu \vdash n$, p. 91
$\mathrm{GL}_n(\mathbb{K})$	matrices carrées : $\mathscr{M}_n(\mathbb{K})$	$P \cdot A = PAP^{-1}$	$(C_{F_1}, C_{F_2}, \ldots, C_{F_s})$, p. 107
$\mathrm{GL}_n(\mathbb{C})$	matrices symétriques : $\mathscr{S}_n(\mathbb{C})$	$P \cdot A = PA\,{}^tP$	$\mathrm{I}_{n,n,r}$, p. 151
$\mathrm{GL}_n(\mathbb{R})$	matrices symétriques : $\mathscr{S}_n(\mathbb{R})$	$P \cdot A = PA\,{}^tP$	$\mathrm{I}_{(p,q)}$, p. 151
$\mathrm{GL}_n(\mathbb{F}_q)$	matrices symétriques inv. : $\mathscr{S}_n(\mathbb{F}_q) \cap \mathrm{GL}_n(\mathbb{F}_q)$	$P \cdot A = PA\,{}^tP$	I_ς, p. 151
$\mathrm{GL}_n(\mathbb{K})$	matrices antisymétriques : $\mathscr{A}_{2n}(\mathbb{K})$	$P \cdot A = PA\,{}^tP$	$J_{n,p}$, p. 160
$\mathrm{PGL}_2(\mathbb{K})$	triplets de points distincts de $\mathbb{P}^1(\mathbb{K})$	action diagonale	$(\infty, 0, 1)$, p. 296

[9]Ou ensemble de paramètres.

Notons, que la notion de forme normale peut aussi concerner des objets géométriques : par exemple un triangle équilatéral peut être vu comme une forme normale pour l'orbite d'un triangle par le groupe affine $\mathrm{GA}_2(\mathbb{R})$, un couple de droites parallèles peut être vu comme une forme normale pour l'orbite d'un couple de droites par le groupe projectif $\mathrm{PGL}_3(\mathbb{R})$! Voir le chapitre XI.

Le tableau qui suit montre le lien fort entre la notion de bases ou de repères en géométrie et l'action simplement transitive d'un groupe.

Par exemple, on sait qu'un *repère affine* de $\mathbb{A}^n$ est par définition un ensemble de $n+1$ points de $\mathbb{A}^n$ tels qu'aucun hyperplan affine ne les contienne tous. Il est bien connu qu'une fois donnés deux repères affines, il existe un unique élément du groupe affine $\mathrm{GA}_n(\mathbb{K})$, *i.e.* le groupe des applications affines bijectives de $\mathbb{A}^n$, qui envoie l'un sur l'autre. Autrement dit, le groupe $\mathrm{GA}_n(\mathbb{K})$ agit de façon simplement transitive sur l'ensemble des repères de $\mathbb{A}^n$.

Actions simplement transitives
$\mathrm{GL}_n(\mathbb{K})$ sur les bases vectorielles de $\mathbb{K}^n$
$\mathrm{O}(n)$ sur les bases orthonormales euclidiennes
$\mathrm{SO}(n)$ sur les bases orthonormales directes euclidiennes
$\mathrm{U}(n)$ sur les bases orthonormales hermitiennes
$\mathrm{SU}(n)$ sur les bases orthonormales directes hermitiennes
$\mathrm{Sp}(n,\mathbb{K})$ sur les bases symplectiques, ex. V-D.17
$\mathrm{Sp}_n(\mathbb{K})$ sur les bases orthonormées symplectiques, ex. V-D.17
$\mathrm{GA}_n(\mathbb{K})$ sur les repères affines, cf p. 331
$\mathrm{PGL}_n(\mathbb{K})$ sur les repères projectifs p. 336

En fait, on peut prendre le problème à l'envers, c'est-à-dire définir la notion de *repère* par la simple transitivité : par repère pour une géométrie, on entendra un moyen d'assembler des points en partie de sorte que le groupe définissant la dite géométrie agisse de façon simplement transitive.

Notons qu'une ligne manque au (dernier) tableau, il s'agit de celle correspondant au groupe d'isotropie $\mathrm{O}(q)$ d'une forme quadratique q. Nous verrons en fait un outil de remplacement : le théorème de Witt qui prolonge les isométries et qui peut être vu comme un théorème de multi-transitivité sous des conditions métriques (cf. p. 159).

F. Exercices

Groupes topologiques : propriétés générales

Dans les exercices qui suivent, G désigne un groupe topologique séparé de neutre e.

F.1. Exercice

1. *Soit $g \in G$. Montrer que les applications $L_g : G \to G$, $h \mapsto gh$ et $R_g : G \to G$, $h \mapsto hg$ sont des homéomorphismes.*
2. *Montrer que si U est un ouvert de G et V une partie quelconque de G, alors les parties $U^{-1} = \{u^{-1}, u \in U\}$ et $UV = \{uv, u \in U, v \in V\}$ sont ouvertes.*
3. *Montrer que si U et V sont compacts, alors UV est compact.*

   En revanche, le produit de deux fermés n'est pas nécessairement fermé : donner un exemple de groupe et de parties fermées U et V telles que UV n'est pas fermé. (Voir l'exercice F.8, *question* 4.*)*

F.2. Exercice

1. *Montrer que l'ensemble des Vx_0 (resp. x_0V), lorsque V parcourt l'ensemble des voisinages de e, est un système fondamental de voisinages du point x_0 de G, c'est-à-dire que tout voisinage de x_0 contient un ensemble de cette forme.*
2. *Soit $f : G \times G \to G$, $(x, y) \mapsto xx_0y^{-1}$. Montrer que f est continue. En déduire que l'ensemble des Vx_0V^{-1} lorsque V parcourt l'ensemble des voisinages de e est un système fondamental de voisinages du point x_0 de G.*

   Soit W un voisinage de x_0, alors $f^{-1}(W)$ est un voisinage de (e, e) ; on peut trouver donc un voisinage ouvert V de e tel que $V \times V \subset f^{-1}(x_0)$. On applique ensuite f.

F.3. Exercice (Propriétés de l'adhérence dans un groupe topologique).

On désigne par $\overline{A}$ l'adhérence de la partie A de G.

1.(a) *Soit U un voisinage de e, x, y dans G. Montrer qu'il existe un ouvert $V \subset U$ tel que $xVyV \subset xyU$.*

   On utilise la multiplication μ de $G \times G$ dans G. La partie $\mu^{-1}(xyU)$ est un ouvert qui contient un ouvert de la forme $xV \times yV$, d'après l'exercice F.2.

   (b) *En déduire que si x et y sont dans les adhérences $\overline{A}$ et $\overline{B}$ de parties A et B de G, alors $xyU \cap AB \neq \emptyset$.*

   Il contient $xVyV \cap AB$, qui est non vide.

   (c) *Montrer que $\overline{A} \cdot \overline{B} \subset \overline{AB}$.*

2. *Montrer que* $\overline{A}^{-1} = \overline{A^{-1}}$.

   Il s'agit d'une double inclusion basée sur le fait que si f est continue, alors $f(\overline{A}) \subset \overline{f(A)}$.

3. *Montrer que pour tout* x, y *dans* G, $x\,\overline{A}\,y = \overline{xAy}$.

   C'est la même chose avec $z \mapsto xzy$ qui est un homéomorphisme.

4. *On suppose que* $ab = ba$ *pour tout couple* (a,b) *de* $A \times B$. *Montrer, en considérant l'application* $f : G \times G \to G$, $(x,y) \mapsto xyx^{-1}y^{-1}$, *que* $ab = ba$ *pour tout couple* (a,b) *de* $\overline{A} \times \overline{B}$.

   $f(A \times B) = \{e\}$ et donc $f(\overline{A} \times \overline{B}) = f(\overline{A \times B})$ est dans $\bar{e} = \{e\}$, car G est séparé.

5. *Montrer que l'adhérence d'un sous groupe* H *de* G *est un sous groupe de* G. *Montrer que* $\overline{H}$ *est abélien si et seulement si* H *l'est.*

F.4. Exercice. *Soit* G *un groupe topologique agissant continûment sur un espace topologique* X.
1) Soit g *dans* G. *Montrer que l'action de* g *est un homéomorphisme de* X.
2) Soit $\mathcal{O}$ *une orbite et* $\overline{\mathcal{O}}$ *son adhérence. Montrer que* $g \cdot \overline{\mathcal{O}} = \overline{g \cdot \mathcal{O}} = \overline{\mathcal{O}}$.
3) En déduire que l'action de G *sur* X *induit une action de* G *sur* $\overline{\mathcal{O}}$, *puis que* $\overline{\mathcal{O}}$ *est une réunion d'orbites pour* G.

F.5. Exercice (Groupe dérivé topologique). *Soit* X *une partie connexe de* G. *On note* H_X *le sous-groupe engendré par* X, *c'est-à-dire :*

$$H_X = \bigcup_{n \geqslant 1} (X \cup X^{-1})^n.$$

1. *Montrer que si* X *contient le neutre de* G, *alors* H_X *est connexe.*

   Que sait-on sur la réunion de connexes qui ont un élément en commun ? Relire l'annexe sur la connexité !

2. *Soit* X_1, X_2 *deux parties connexes de* G *contenant l'identité et*

   $$X = \{x_1 x_2 x_1^{-1} x_2^{-1},\ x_1 \in X_1,\ x_2 \in X_2\}.$$

   Montrer que l'adhérence de H_X *est un sous-groupe de* G *et qu'il est connexe.*

   Remarque : si $X_1 = X_2 = G$, ce sous-groupe est appelé le *sous-groupe dérivé topologique* de G.

F.6. Exercice (Composante connexe du neutre). *Soit G_0 la composante neutre de G, c'est-à-dire la composante connexe qui contient le neutre.*

1. *Montrer que G_0 est un sous-groupe fermée.*

   Remarque : G_0 peut ne pas être ouvert. Voir le contre-exemple dans l'exercice F.8.

   Si G_0 est connexe, alors $\overline{G_0}$ l'est aussi.

2. *Montrer que la composante connexe d'un élément g est gG_0.*

   Un homéomorphisme envoie une composante connexe sur une composante connexe.

3. *Montrer que G_0 est stable par conjugaison.*

   $xG_0 = G_0x$ puisque ce sont tout deux la composante connexe de x.

4. *Montrer que si H est un sous-groupe ouvert inclus dans G_0, alors $H = G_0$.*

   Si H est un sous groupe ouvert, alors il est fermé, car son complémentaire est l'ouvert $G_0 \backslash H = \bigcup_{g \notin H} gH$. Comme il est ouvert et fermé dans une composante connexe, on a égalité.

5. *Montrer que si G est connexe, alors pour tout voisinage V du neutre, on a $G = H_V$, où la notation H_V est définie dans l'exercice* F.5.

   C'est la question précédente.

6. *Montrer que G_0 est ouverte si et seulement si le neutre admet un voisinage connexe.*

   L'implication est claire, on peut prendre pour l'ouvert G_0. La réciproque vient de la question qui précède.

Groupes topologiques : exemples

F.7. Exercice (Sous-groupes à un paramètre du groupe linéaire). *Soit n un entier naturel et $f : \mathbb{R} \to \mathrm{GL}_n(\mathbb{C})$ un morphisme de groupes continu.*

1. *Soit $\alpha > 0$. Vérifier que l'on a :*
   $\left(\int_0^\alpha f(s)\mathrm{d}s\right) f(t) = \int_t^{t+\alpha} f(s)\mathrm{d}s$ *pour tout t réel.*
2. *Montrer que si α est assez petit, la valeur moyenne de f sur $[0, \alpha]$,* i.e. $\frac{1}{\alpha}\int_0^\alpha f(s)\mathrm{d}s$, *est une matrice inversible.*
3. *En déduire que f est dérivable sur $\mathbb{R}$ et que l'on a :*
   $f'(t) = f'(0)f(t)$ *pour tout réel t.*
4. *Conclure : $f(t) = \exp tf'(0)$ pour tout réel t.*

Remarque : on observe ici que l'association d'une hypothèse algébrique (morphisme de groupe) et d'une hypothèse topologique (continuité) est très contraignante : les seuls morphismes continus sont indéfiniment différentiables et même analytiques. On retrouvera ce phénomène avec le théorème de Cartan au chapitre IX.

F.8. Exercice (Sous-groupes de $\mathbb{R}$ et merveilleuses applications)

1. *Soit H un sous-groupe non trivial de $\mathbb{R}$. Soit $a = \inf(H \cap]0, +\infty[)$.*
   (a) Montrer que si a est non nul, alors $H = a\mathbb{Z}$.
   (b) Montrer que si a est nul, alors H est dense dans $\mathbb{R}$.
   (c) En déduire que tout sous-groupe de $\mathbb{R}$ est soit cyclique, soit dense.
2. *On veut étudier le groupe $H = \mathbb{Z} + \alpha\mathbb{Z}$, où α est un réel.*
   (a) On suppose que $\alpha = u/v \in \mathbb{Q}$, $(u, v) \in \mathbb{Z} \times \mathbb{Z}^*$. Montrer que $H = d/v\mathbb{Z}$, où $d = \operatorname{pgcd}(u, v)$.
   (b) On suppose α irrationnel. Montrer que H est dense dans $\mathbb{R}$.
   (c) En déduire que tout élément de $[-1, 1]$ est valeur d'adhérence de la suite $(\cos n)_{n\in\mathbb{N}}$.
   Le nombre 2π est irrationnel ; le sous-groupe $\{m + 2\pi n, m, n \in \mathbb{Z}\}$ est de ce fait dense dans $\mathbb{R}$. Prendre ensuite le cosinus.
   (d) En déduire par ailleurs qu'il existe une infinité de puissances de 2 dont l'écriture décimale commence par 9.
   Il faut montrer qu'il existe une infinité de couples d'entiers (m, n) tels que
   $$9\times10^m \leqslant 2^n < 10\times10^m.$$
   Prendre le logarithme de base 10 et utiliser le fait que $\dfrac{\log(2)}{\log(9)}$ est irrationnel pour montrer que le groupe $\left\{n\dfrac{\log(2)}{\log(9)} + m,\ n, m \in \mathbb{Z}\right\}$ est dense dans $\mathbb{R}$.
   (e) On considère le groupe G, quotient de $\mathbb{R}^2$ par $\mathbb{Z}^2$ muni de la topologie quotient. On note $\pi : \mathbb{R}^2 \to G$ la projection naturelle. Pour θ irrationnel fixé, on note H_0 le sous-groupe de $\mathbb{R}^2$ formé des couples (x, y) satisfaisant à $y = \theta x$, H son image par la projection π et $\tilde{H}$ la saturation de H_0, c'est-à-dire : $\tilde{H} = \pi^{-1}\big(\pi(H_0)\big)$. Démontrer que $\tilde{H}$ est dense dans $\mathbb{R}^2$; en déduire que H est dense dans G et étudier la séparation dans le groupe quotient G/H.
3. *Prouver que le groupe $\mathbb{R}/2\pi\mathbb{Z}$ (le cercle, quoi !) contient exactement un sous-groupe fini de tout cardinal et que ses sous-groupes infinis sont denses.*

   On pourra étudier l'image réciproque d'un sous-groupe par la projection naturelle $\mathbb{R} \to \mathbb{R}/2\pi\mathbb{Z}$.

4. On considère le sous-groupe $\mathbb{Q}$ de $\mathbb{R}$ muni de la topologie métrique induite.

(a) Montrer que $\mathbb{Q}$ n'est pas discret.

(b) Montrer que la composante connexe de l'élément neutre 0 est $\{0\}$.

F.9. Exercice (Sous-groupes fermés et réseaux de $\mathbb{R}^n$)

Soit n un entier naturel non nul et Γ un sous-groupe fermé de $\mathbb{R}^n$. On fixe une norme $||\cdot||$ sur $\mathbb{R}^n$ et on note 0 le neutre de $\mathbb{R}^n$ et Γ.

1. On suppose que 0 est un point isolé de Γ et que le sous-espace vectoriel engendré par Γ est $\mathbb{R}^n$. Soit $(e_1, \ldots, e_n)$ une base de $\mathbb{R}^n$ contenue dans Γ et

$$P = \Big\{ \sum_{i=1}^{n} \lambda_i e_i, \quad \forall i,\ 0 \leqslant \lambda_i \leqslant 1 \Big\}.$$

(a) Démontrer que $\Gamma \cap P$ est un ensemble fini.

Saviez-vous que toute partie discrète et compacte est finie ?

(b) Démontrer que le groupe Γ est engendré par $\mathscr{B} \cup (\Gamma \cap P)$.

On admet que tout groupe abélien de type fini sans torsion est isomorphe à $\mathbb{Z}^r$ pour un entier r bien déterminé. Ici, on a nécessairement $r = n$.

2. On suppose que 0 est un point d'accumulation de Γ. Soit $(\gamma_p)_{p\in\mathbb{N}}$ une suite de points de $\Gamma \setminus \{0\}$ qui converge vers 0. Pour $p \in \mathbb{N}$, on note

$$k_p = \max\{k \in \mathbb{N}, \quad ||k\gamma_p|| \leqslant 1\}.$$

La suite $(k_p\gamma_p)_{p\in\mathbb{N}}$ est bornée. Quitte à extraire une sous-suite, on suppose qu'elle converge vers un élément $\gamma \in \Gamma$.

(a) Vérifier que l'on a : $||\gamma|| = 1$.

(b) Soit $t \in \mathbb{R}$. On pose pour tout entier p : $m_p = \lfloor tk_p \rfloor$, où $\lfloor\cdot\rfloor$ désigne la partie entière. Quelle est la limite de la suite $(m_p\gamma_p)_{p\in\mathbb{N}}$?

(c) En déduire que $\mathbb{R}\gamma \subset \Gamma$.

3. On suppose que 0 est un point d'accumulation de Γ. Soit F l'ensemble des éléments γ de Γ tels que $\mathbb{R}\gamma \subset \Gamma$.

(a) Vérifier que F est un sous-espace vectoriel de $\mathbb{R}^n$.

(b) Soit F' un supplémentaire de F dans $\mathbb{R}^n$ et $\Gamma' = \Gamma \cap F'$. Montrer que 0 n'est pas un point d'accumulation de Γ' et que l'on a l'isomorphisme de groupes :

$$\Gamma \simeq F \times \Gamma'.$$

4. *On suppose que* Γ *est un réseau de* $\mathbb{R}^n$*, c'est-à-dire un sous-groupe discret tel que le quotient* $\mathbb{R}^n/\Gamma$ *a une mesure finie (pour la mesure image de la mesure de Lebesgue par la projection canonique). Prouver que* Γ *est le groupe engendré par une base de* $\mathbb{R}^n$*.*

F.10. Exercice. *Montrer que* $(\mathbb{C}^n, +)$ *est isomorphe à un sous-groupe de* $\mathrm{GL}_{n+1}(\mathbb{C})$*.*

Commencer par le cas $n = 1$ et considérer l'ensemble $\left\{ \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}, a \in \mathbb{C} \right\}$.

F.11. Exercice (Surjectivité de l'exponentielle complexe). *Soit* exp : $\mathbb{C} \to \mathbb{C}^*$ *l'exponentielle complexe.*

1. *Prouver que l'image de l'exponentielle contient un voisinage de* 1.

   Noter que la différentielle en 0 est bijective.

2. *Avec l'exercice* F.6*, question 5., prouver que l'exponentielle est surjective.*

3. *Prouver que l'exponentielle est une application ouverte.*

   Utiliser le principe de translation.

F.12. Exercice (L'exponentielle complexe (bis)). *On considère l'application exponentielle*

$$\mathrm{Exp} : \mathbb{C} \to \mathbb{C}^*, \quad z \mapsto e^{2i\pi z}.$$

1. *Montrer que* Exp *est un homomorphisme de groupes topologiques et déterminer son noyau.*

   C'est $\mathbb{Z}$, bien sûr.

2. *On considère le pavé ouvert*

$$P = \{z = x + iy \mid x_0 < x < x_1,\, y_0 < y < y_0 + 1\},$$

   où $x_0 < x_1$*,* y_0 *sont des réels. Montrer que la restriction de* Exp *à* P *est un homéomorphisme sur son image.*

   C'est clairement continu et bijectif, puis un argument de compacité prouve qu'un fermé s'envoie sur un fermé et donc un ouvert s'envoie vers un ouvert.

3. *En déduire que pour chaque point* w *de* $\mathbb{C}^*$*, il existe un voisinage ouvert* W *tel que* $\mathrm{Exp}^{-1}(W)$ *est réunion disjointe d'ouverts connexes* P_n *et que pour chaque* n*,* Exp *réalise un homéomorphisme de* P_n *dans* W*.*

Actions de groupes et théorème d'homéomorphisme

F.13. Exercice. *Soit G un groupe topologique agissant sur un ensemble X. Soit x et y deux points de X dans une même orbite sous G. Vérifier que les espaces G/G_x et G/G_y, munis de la topologie quotient, sont homéomorphes.*

F.14. Exercice (Compacité de groupes classiques). *Soit n un entier naturel. Montrer que groupes classiques $\mathrm{O}(n,\mathbb{R})$, $\mathrm{U}_n(\mathbb{C})$, $\mathrm{SO}(n,\mathbb{R})$, $\mathrm{SU}_n(\mathbb{C})$ sont compacts.*

On définit le groupe symplectique complexe comme le stabilisateur d'une forme bilinéaire antisymétrique non dégénérée (voir section V-4) :

$$\mathrm{Sp}_{2n}(\mathbb{C}) = \left\{A \in \mathrm{GL}_{2n}(\mathbb{C}) \ , \ {}^tAJ_nA = J_n\right\}, \quad \text{où } J_n = \begin{pmatrix} 0 & -\mathrm{I}_n \\ \mathrm{I}_n & 0 \end{pmatrix}.$$

On définit le groupe symplectique compact par $\mathrm{Sp}(n,\mathbb{R}) = \mathrm{Sp}_{2n}(\mathbb{C}) \cap \mathrm{U}_{2n}(\mathbb{C})$. Prouver que le groupe symplectique compact est bien compact.

F.15. Exercice (Connexité du groupe spécial linéaire complexe). *Soit n un entier naturel. Montrer que le groupe $\mathrm{SL}_n(\mathbb{C})$ est connexe.*

Procéder par récurrence en utilisant l'action naturelle sur $\mathbb{C}^n \setminus \{0\}$.

F.16. Exercice (Connexité du groupe spécial orthogonal). *Soit n un entier positif. On rappelle que la sphère unité $\mathbf{S}_{n-1}$ de $\mathbb{R}^n$ est connexe par arcs. On considère l'action naturelle*

$$\alpha : \mathrm{SO}(n,\mathbb{R}) \times \mathbb{R}^n \longrightarrow \mathbb{R}^n, \quad (A,x) \longmapsto Ax.$$

1. *Montrer que cette action se restreint en une action transitive de $\mathrm{SO}(n,\mathbb{R})$ sur $\mathbf{S}_{n-1}$.*
2. *Calculer le groupe d'isotropie (ou stabilisateur) du point $N = (1,0,\cdots,0)$.*
3. *Montrer que $\mathrm{SO}(n,\mathbb{R})/\mathrm{SO}(n-1,\mathbb{R})$ est un espace topologique compact homéomorphe à $\mathbf{S}_{n-1}$.*
4. *Montrer par récurrence que $\mathrm{SO}(n,\mathbb{R})$ est connexe.*

F.17. Exercice (Connexité des groupes unitaires). *On considère l'opération*

$$\alpha : \mathrm{SU}_n(\mathbb{C}) \times \mathbb{C}^n \longrightarrow \mathbb{C}^n, \quad (A,x) \longmapsto Ax.$$

1. *Montrer que cette action se restreint en une action transitive de $\mathrm{SU}_n(\mathbb{C})$ sur la sphère unité $\mathbf{S}_{2n-1}$ de $\mathbb{C}^n$.*
2. *Calculer le sous-groupe d'isotropie du point $N = (1,0,\cdots,0)$.*

3. *Montrer que le quotient* $\mathrm{SU}_n(\mathbb{C})/\mathrm{SU}_{n-1}(\mathbb{C})$ *est un espace compact homéomorphe à* $\mathbf{S}_{2n-1}$.

4. *Montrer que par récurrence que* $\mathrm{SU}_n(\mathbb{C})$ *est connexe.*

5. *Montrer que* $\mathrm{U}_n(\mathbb{C})/\mathrm{SU}_n(\mathbb{C})$ *est homéomorphe à* $\mathbf{S}_1$. *En déduire que le groupe unitaire* $\mathrm{U}_n(\mathbb{C})$ *est connexe.*

F.18. Exercice (Homographies du demi-plan de Poincaré). *Soit* $\mathscr{H}$ *le demi-plan de Poincaré :*

$$\mathscr{H} = \{z \in \mathbb{C} \ , \ \ \mathrm{Im}(z) > 0\}.$$

1. *Montrer que l'application*

$$\alpha : \mathrm{SL}_2(\mathbb{R}) \times \mathscr{H} \to \mathscr{H}, \quad \left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, z\right) \mapsto \frac{az+b}{cz+d}$$

est bien définie.

2. *Montrer que* α *est une action continue de* $\mathrm{SL}_2(\mathbb{R})$ *sur* $\mathscr{H}$. *Quel est son noyau ? Calculer le groupe d'isotropie de* i.

3. *Montrer que* $\mathscr{H}$ *est homéomorphe au quotient* $\mathrm{SL}_2(\mathbb{R})/\mathrm{SO}_2(\mathbb{R})$.

F.19. Exercice (Structures complexes sur un espace vectoriel réel). *Soit* p *un entier positif. On note*

$$\mathscr{C} = \{C \in \mathscr{M}_p(\mathbb{R}) \ , \ \ C^2 = -\mathrm{I}_p\}$$

et on appelle structures complexes de $\mathbb{R}^p$ *les éléments de* $\mathscr{C}$. *Comme partie d'un espace de matrices,* $\mathscr{C}$ *est muni de la topologie naturelle.*

1. *Montrer que* $\mathbb{R}^p$ *admet une structure complexe si et seulement si* p *est pair.*

2. *Dans cette question, on suppose que* $p = 2$.

   (a) *Montrer que la trace d'une structure complexe est nulle et dessiner l'image réciproque de* $\mathscr{C}$ *par l'application*

$$\mathbb{R}^3 \longrightarrow \mathrm{Ker}\,\mathrm{tr}, \quad (x,y,z) \longmapsto \begin{pmatrix} x & y \\ z & -x \end{pmatrix}.$$

   (b) *Est-ce que* $\mathscr{C}$ *est connexe ? Comment interpréter ce fait ?*

   (c) *Soit* (x,y) *un élément du demi-plan de Poincaré* $\mathscr{H}$ *(exercice* F.18*). Montrer qu'il existe un unique élément de* $\mathscr{C}$ *dont la première colonne est* ${}^t(x\ y)$. *Prouver que l'application* $\mathscr{H} \to \mathscr{C}$ *ainsi construite est un homéomorphisme de* $\mathscr{H}$ *sur son image dans* $\mathscr{C}$, *que l'on caractérisera.*

3. *On suppose que p est pair et on écrit $p = 2n$. Soit C une structure complexe sur $\mathbb{R}^{2n}$. Montrer que l'application*

$$\mathbb{C} \times \mathbb{R}^{2n} \to \mathbb{R}^{2n}, \quad (a + ib, X) \to aX + bCX$$

définit une structure de $\mathbb{C}$-espace vectoriel sur $\mathbb{R}^{2n}$. On la désignera comme étant « associée à C »[10].

4. *Soit $(e_1, \ldots, e_d)$ une famille finie de vecteurs de $\mathbb{R}^{2n}$. Montrer que $(e_1, \ldots, e_d)$ est une base du $\mathbb{C}$-espace vectoriel associé à C si et seulement si $(e_1, \ldots, e_d, Ce_1, \ldots, Ce_d)$ est une base du $\mathbb{R}$-espace vectoriel $\mathbb{R}^{2n}$. En déduire la dimension du $\mathbb{C}$-espace vectoriel associé à C.*

5. *Vérifier que l'application $\varphi_C : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$, $v \mapsto Cv$ est une application linéaire pour les deux structures d'espace vectoriel. En déduire que $\mathscr{C}$ est la classe de similitude de la forme normale*

$$J_n = \begin{pmatrix} 0_n & -\mathrm{I}_n \\ \mathrm{I}_n & 0_n \end{pmatrix}.$$

Déterminer la matrice de φ_C dans une base de l'espace vectoriel associé à C.

6. *Utiliser la matrice J_n pour donner un plongement naturel de $\mathrm{GL}_n(\mathbb{C})$ dans $\mathrm{GL}_{2n}(\mathbb{R})$.*

7. *Montrer que $\mathscr{C}$ est homéomorphe à $\mathrm{GL}_{2n}(\mathbb{R})/\mathrm{GL}_n(\mathbb{C})$.*

8. *(Question subsidiaire.) Trouver un automorphisme θ de $\mathrm{GL}_{2n}(\mathbb{R})$ dont $\mathrm{GL}_n(\mathbb{C})$ est le groupe des points fixes.*

   Penser à la conjugaison par μ_i, où μ_i est la multiplication par i dans $\mathbb{C}^n \simeq \mathbb{R}^{2n}$.

On a ainsi une topologie sur l'ensemble des structures complexes, premier pas vers une étude géométrique de cet objet algébrique.

F.20. Exercice (Grassmanniennes complexes). *Soit n un entier naturel non nul et d un entier compris entre 1 et $n-1$. Soit*

$$P_d = \left\{ \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}, \ A \in \mathrm{GL}_d(\mathbb{C}),\ D \in \mathrm{GL}_{n-d}(\mathbb{C}),\ B \in \mathscr{M}_{d,n-d}(\mathbb{C}) \right\}.$$

1. *Identifier $\mathrm{GL}_n(\mathbb{C})/P_d$ à la grassmannienne $\mathrm{Gr}_{d,n}(\mathbb{C})$ des d-plans de $\mathbb{C}^n$.*

2. *Prouver que l'action du groupe spécial unitaire $\mathrm{U}_n(\mathbb{C})$ sur $\mathrm{Gr}_{d,n}(\mathbb{C})$ est transitive. En déduire que $\mathrm{Gr}_{d,n}(\mathbb{C})$ est un espace compact.*

[10] Attention, on a donc un seul ensemble, $\mathbb{R}^{2n}$, mais deux structures d'espaces vectoriels sur des corps différents.

F.21. Exercice (Topologie sur les espaces de drapeaux). *Soit n un entier naturel non nul et $I = (p_1, \dots, p_r)$ une suite finie d'entiers naturels non nuls dont la somme vaut n,* i.e. *: $\sum_k p_k = n$. On appelle* drapeau de type $I = (p_1, p_2, \dots, p_r)$ *toute famille $F_I = (F_1, F_2, \cdots, F_r)$ de sous-espaces vectoriels de $\mathbb{C}^n$ tels que :*

$$F_1 \subset F_2 \subset \cdots \subset F_r \quad \text{et} \quad \forall k \in \{1, \dots, r\},\ \dim F_k = p_1 + \cdots + p_k$$

On désigne par $\mathscr{F}_I$ l'ensemble des drapeaux de type I (appelé aussi variété de drapeaux de type I, car il s'agit d'une variété algébrique[11]*).*

1. *Reconnaître l'espace $\mathscr{F}_{(1,n-1)}$.*
2. *Montrer que $\mathrm{GL}_n(\mathbb{C})$ agit transitivement sur $\mathscr{F}_I$.*
3. *Le I-drapeau standard est le drapeau dont le k^{e} espace F_k est engendré par les $\dim(F_k)$ premiers vecteurs de la base canonique de $\mathbb{C}^n$ pour tout k. Décrire le stabilisateur du I-drapeau standard (ce stabilisateur est un exemple de* sous-groupe parabolique *de $\mathrm{GL}_n(\mathbb{C})$).*
4. *En déduire une topologie sur $\mathscr{F}_I$.*
5. *Prouver que l'action de $\mathrm{SU}_n(\mathbb{C})$ sur $\mathscr{F}_I$ obtenue par restriction de l'action de $\mathrm{GL}_n(\mathbb{C})$ est toujours transitive. En déduire que $\mathscr{F}_I$ est un espace compact.*

F.22. Exercice (Produits scalaires euclidiens). *Soit n un entier strictement positif et $\mathscr{S}_n^{++}$ l'ensemble des matrices réelles $n \times n$ symétriques définies positives.*

1. *Montrer que l'action de $\mathrm{GL}_n(\mathbb{R})$ par congruence sur $\mathscr{S}_n^{++}$ permet de réaliser $\mathscr{S}_n^{++}$ comme quotient de $\mathrm{GL}_n(\mathbb{R})$ par $\mathrm{O}(n, \mathbb{R})$.*
2. *On suppose que $n = 2$.*
   (a) *Dessiner $\mathscr{S}_2^{++}$.*
   (b) *À un élément $z = (x, y)$ du demi-plan de Poincaré (*F.18*), on associe le produit scalaire pour lequel $\big((1,0),(x,y)\big)$ est une base orthonormée. Montrer que l'application $f : \mathscr{H} \to \mathscr{S}_2^{++}$ ainsi définie est équivariante,* i.e. *commute à l'action de $\mathrm{SL}_2(\mathbb{R})$. Dessiner l'image de f.*
   (c) *On fait agir $\mathbb{R}_+^*$ par multiplication sur $\mathscr{S}_2^{++}$ et on munit $\mathscr{S}_2^{++}/\mathbb{R}_+^*$ de la topologie quotient* (D)*. Montrer que la composée de f et de la projection canonique est un homéomorphisme de $\mathscr{H}$ sur $\mathscr{S}_2^{++}/\mathbb{R}_+^*$. Décrire sa réciproque.*

   Étant donné un élément $\mathbf{A}$ de $\mathscr{S}_2^{++}/\mathbb{R}_+^*$, choisir la matrice A de $\mathbf{A}$ telle que $(1\ 0)A^{\mathrm{t}}(1\ 0) = 1$; compléter $(1,0)$ en une base orthonormée directe pour A.
   (d) *Établir un lien direct entre $\mathscr{S}_2^{++}/\mathbb{R}_+^*$ et l'espace des structures complexes $\mathscr{C}$ (exercice* F.19*).*

[11] Ceci s'éclaircira un peu au §I-2 du second tome.

F.23. Exercice (Trois topologies sur l'espace projectif ?). *Soit n un entier naturel non nul. L'espace projectif réel $\mathbb{P}^{n-1}$ de dimension $n-1$ est l'ensemble des droites de $\mathbb{R}^n$; il possède trois réalisations comme ensemble :*

(i) l'application naturelle $(\mathbb{R}^n \setminus \{0\}) \to \mathbb{P}^{n-1}$, $v \mapsto \mathbb{R}v$ est constante sur les orbites sous $\mathbb{R}^*$ et induit une bijection $(\mathbb{R}^n \setminus \{0\})/\mathbb{R}^* \to \mathbb{P}^{n-1}$, $\mathbb{R}^*v \mapsto \mathbb{R}v$; l'espace projectif hérite de la topologie définie au § D ;

(ii) l'action transitive de $\mathrm{GL}_n(\mathbb{R})$ sur les droites induit, via le choix d'un vecteur non nul, disons le premier vecteur de base e_1, une bijection $\overline{f} : \mathrm{GL}_n(\mathbb{R})/P \to \mathbb{P}^{n-1}$, $gP \mapsto g\mathbb{R}e_1$; l'espace projectif hérite de la topologie de $\mathrm{GL}_n(\mathbb{R})/P$;

(iii) de même, l'action de $\mathrm{SO}(n,\mathbb{R})$ induit une bijection $\mathrm{SO}(n,\mathbb{R})/K \to \mathbb{P}^{n-1}$, $kK \mapsto k\mathbb{R}e_1$, où K est le groupe d'isotropie de $\mathbb{R}e_1$.

On a déjà vu au § 4.3 que les topologies des réalisations (ii) et (iii) coïncident. On considère que $\mathbb{P}^{n-1}$ est muni de la topologie donnée par la bijection (i) et on veut prouver que $\overline{f}$ est un homéomorphisme.

1. *Prouver que l'application $f : \mathrm{GL}_n(\mathbb{R}) \longrightarrow \mathbb{R}^n \setminus \{0\}$, $g \longmapsto ge_1$ est ouverte.*
2. *En déduire que l'application $\overline{f}$ induite par f par passage au quotient est ouverte, puis que c'est un homéomorphisme.*

F.24. Exercice (Classes de similitude des réseaux de $\mathbb{C}$). *Soit $\mathscr{R}$ l'ensemble des réseaux de l'espace vectoriel réel $\mathbb{C}$ (c'est-à-dire des sous-groupes isomorphes à $\mathbb{Z}^2$ non contenus dans une droite réelle, voir l'exercice F.9). L'action*[12] *par multiplication de $\mathbb{C}^*$ sur $\mathbb{C}$ induit une action de $\mathscr{R}$. On s'intéresse*[13] *au quotient $\mathscr{R}/\mathbb{C}^*$ des réseaux à homothétie près, qu'on va réaliser comme un double quotient de $\mathrm{SL}_2(\mathbb{R})$.*

Suivant [34], chapitre VII, § 2.2, on note M l'ensemble des couples $(\omega_1, \omega_2) \in \mathbb{C}^* \times \mathbb{C}^*$ tels que $\mathrm{Im}\,\omega_1/\omega_2 > 0$.

1. *Prouver que l'application*

$$\pi : M \longrightarrow \mathscr{H}, \quad (\omega_1, \omega_2) \longmapsto \omega_1/\omega_2,$$

où $\mathscr{H}$ est le demi-plan de Poincaré (exercice F.18), est surjective et induit par passage au quotient une bijection $\overline{\pi} : M/\mathbb{C}^* \to \mathscr{H}$.

[12] Ce sont des similitudes du plan affine euclidien $\mathbb{C}$.

[13] Si Γ est un réseau de $\mathbb{C}$, le quotient $\mathbb{C}/\Gamma$ est une courbe elliptique, au sens suivant. Comme groupe et comme espace topologique (ou même comme variété différentiable), $\mathbb{C}/\Gamma$ s'identifie à $(\mathbb{R}/\mathbb{Z})^2$ – un tore. Mais on peut le voir comme une courbe complexe – par exemple, on peut le paramétrer par une fonction méromorphe d'une variable complexe. Voir, par exemple, la remarque VI.2.7 de [5] ou le paragraphe VIII.5 de [8]. Or, on peut montrer que deux courbes elliptiques $\mathbb{C}/\Gamma$ et $\mathbb{C}/\Gamma'$ sont isomorphes (comme groupes et comme courbes complexes) si et seulement si Γ et Γ' sont proportionnels. Ainsi, $\mathscr{R}/\mathbb{C}^*$ s'identifie aux classes d'isomorphisme de courbes elliptiques.

2. *Prouver la surjectivité de l'application*

$$\Lambda : M \longrightarrow \mathscr{R}, \quad (\omega_1, \omega_2) \longmapsto \mathbb{Z}\omega_1 \oplus \mathbb{Z}\omega_2.$$

3. *Étant donné*

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}),$$

et $(\omega_1, \omega_2) \in M$, *on pose*

$$\omega_1' = a\omega_1 + b\omega_2, \quad \omega_2' = c\omega_1 + d\omega_2,$$

Vérifier que (ω_1', ω_2') *est une base de* $\Lambda(\omega_1, \omega_2)$, *et, que si on pose* $z = \omega_1/\omega_2$ *et* $z' = \omega_1'/\omega_2'$, *on a :* $z' = g \cdot z$. *En déduire une action de* $\mathrm{SL}_2(\mathbb{Z})$ *sur* M.

4. *Prouver que deux éléments de* M *définissent le même réseau de* $\mathbb{R}^2$ *si et seulement s'ils sont dans la même orbite sous* $\mathrm{SL}_2(\mathbb{Z})$. *En d'autres termes,* Λ *induit une identification* $M/\mathrm{SL}_2(\mathbb{Z}) \simeq \mathscr{R}$.

5. *En utilisant le fait que les actions de* $\mathbb{C}^*$ *et* $\mathrm{SL}_2(\mathbb{Z})$ *commutent et l'action de* $\mathrm{SL}_2(\mathbb{Z})$ *sur* $\mathscr{H}$, *prouver que l'application* $(\omega_1, \omega_2) \to \omega_1/\omega_2$ *induit une identification :* $\mathscr{R}/\mathbb{C}^* \to \mathscr{H}/\mathrm{SL}_2(\mathbb{Z})$.

6. *Construire une topologie sur l'ensemble* $\mathscr{R}/\mathbb{C}^*$. *Prouver qu'elle fait de l'espace des réseaux un espace séparé mais non compact (ce dernier point est difficile : considérer la suite de réseaux* $\big(\Lambda(1, in)\big)_{n\in\mathbb{N}}$*).*

Produits semi-directs (topologiques)

F.25. Exercice (La gauche et la droite). *Étant donné deux groupes topologiques* H *et* K *et une action continue de* H *sur* K, *définie par un morphisme* $\varphi : H \to \mathrm{Aut}(K)$, $h \mapsto \varphi_h$, *le cours donne une définition du produit semi-direct* $K \rtimes_\varphi H$. *On peut en donner une variante en permutant le groupe de gauche et le groupe de droite. On cherche à définir un groupe topologique ainsi :*
– comme espace topologique sous-jacent, c'est $H \times K$,
– le produit est défini par :

$$\forall (h,k),(h',k'), \quad (h,k)\cdot(h',k') = (hh', \varphi_{h'^{-1}}(k)k').$$

La notation consacrée pour cette version est $H \ltimes_\varphi K$.

Vérifier que $H \ltimes_\varphi K$ *est bien un groupe topologique et que l'application « volte » suivante est un isomorphisme et un homéomorphisme :* $\tau : K \ltimes_\varphi H \to K \rtimes_\varphi H$, $(h,k) \mapsto \big(\varphi_h(k), h\big)$.

Remarque : L'application τ *est parfaitement naturelle si on pense qu'après identification de sous-groupes du produit semi-direct isomorphes à* H *et* K, *l'action de* K *sur* H *s'identifie à la conjugaison dans le produit semi-direct. La formule* $hkh^{-1} = \varphi_h(k)$ *donne* $hk = \varphi_h(k)h$, *ce qui justifie la formule pour* τ.

F.26. Exercice (Groupe affine). *Soit $\mathscr{E}$ un espace affine dirigé par l'espace vectoriel réel E de dimension finie n. Soit $\mathrm{GA}(\mathscr{E})$, ou plus simplement GA, le groupe des bijections affines de $\mathscr{E}$ et $\mathrm{GL}(E)$, ou simplement GL, le groupe des bijections linéaires de E. On souhaite démontrer et interpréter un isomorphisme (non canonique) :*

$$\mathrm{GA}(\mathscr{E}) \simeq E \rtimes \mathrm{GL}(E).$$

Soit T le groupe des translations. Il est isomorphe à E par l'application qui, à $v \in E$, associe la translation $t_v : \mathscr{E} \to \mathscr{E}$ de vecteur v. Soit O un point de $\mathscr{E}$ et GA_O le stabilisateur de O.

1. *Soit $f \in \mathrm{GA}$ et $v \in E$. Donner une expression simple de ft_vf^{-1}. En déduire que T est distingué.*
2. *Vérifier que GA_O est isomorphe à GL. Montrer que toute application $f \in \mathrm{GA}$ s'écrit de façon unique comme une composée $t_v \circ g$ (resp. $g \circ t_v$) avec $v \in E$ et $g \in \mathrm{GA}_O$.*
3. *Étant donné une application $f \in \mathrm{GA}$, on note $\mathrm{d}f \in \mathrm{GL}$ l'application linéaire associée (oui : c'est bien sa différentielle au sens de la géométrie différentielle...). L'utiliser pour définir la suite exacte :*

$$1 \to T \longrightarrow \mathrm{GA}(\mathscr{E}) \xrightarrow{\mathrm{d}} \mathrm{GL}(E) \to 1.$$

4. *Exhiber une section de d et en déduire la structure de produit semi-direct (topologique).*
5. *On suppose à présent que $\mathscr{E}$ est un espace affine euclidien. On note $\mathrm{Is}(\mathscr{E})$ les groupe des isométries (affines) de $\mathscr{E}$ et $\mathrm{Is}(E)$ le groupe des isométries vectorielles de E. Montrer que l'on a un isomorphisme : $\mathrm{Is}(\mathscr{E}) \simeq T \rtimes \mathrm{Is}(E)$.*

F.27. Exercice (Dévissage de groupes classiques). *Soit n un entier naturel non nul.*

1. *Démontrer que $\mathrm{GL}_n(\mathbb{R})$ est isomorphe à un produit semi-direct topologique de $\mathrm{SL}_n(\mathbb{R})$ par $\mathbb{R}^*$. Idem sur $\mathbb{C}$.*
2. *Démontrer que le groupe orthogonal $\mathrm{O}(n,\mathbb{R})$ est isomorphe à un produit semi-direct topologique de $\mathrm{SO}(n,\mathbb{R})$ par $\mathbb{Z}/2\mathbb{Z}$. Pour quelles valeurs de n ce produit est-il direct (c'est-à-dire que l'action de $\mathbb{Z}/2\mathbb{Z}$ sur $\mathrm{SO}(n)$ est triviale) ?*

F.28. Exercice (Stabilisateur d'un sous-espace vectoriel). *Soit n un entier non nul et $d \in \{1,\dots,n\}$. On fait agir GL_n sur l'ensemble des sous-espaces vectoriels de dimension d. Exhiber une structure de produit semi-direct sur le stabilisateur d'un sous-espace.*

F.29. Exercice (Le groupe quaternionique n'est pas un produit semi-direct). *Soit H_8 le groupe (parfois appelé « quaternionique ») :*

$$H_8 = \{\pm 1, \pm i, \pm j, \pm k\},$$

où le produit est défini par :

$$ij = k,\ jk = i,\ ki = j, i^2 = j^2 = k^2 = -1,$$

où 1 est neutre et la multiplication par -1 suit les règles naturelles.

1. *Quel est le centre Z de H_8 ? Quel est le quotient H_8/Z ? Quels sont les ordres des éléments de H_8 ? Quels sont les sous-groupes de H_8 ?*
2. *Est-ce que H_8 est isomorphe à un produit semi-direct non trivial ?*

F.30. Exercice (Groupe de Heisenberg). *Soit n un entier naturel non nul. On note G le groupe de Heisenberg : c'est le sous-groupe de $\mathrm{GL}_{n+2}(\mathbb{R})$ formé des matrices de la forme*

$$[x,y,z] = \begin{pmatrix} 1 & x_1 & \cdots & x_n & z \\ & \ddots & 0 & 0 & y_1 \\ & & \ddots & 0 & \vdots \\ & (0) & & \ddots & y_n \\ & & & & 1 \end{pmatrix},$$

$$x = (x_1, \ldots, x_n) \in \mathbb{R}^n,\ y = (y_1, \ldots, y_n) \in \mathbb{R}^n,\ z \in \mathbb{R}.$$

1. *Calculer le produit $[x,y,z][x',y',z']$ de deux éléments de G.*
2. *Quel est le centre Z de G ? Montrer que le quotient G/Z est isomorphe à $\mathbb{R}^{2n}$ comme groupes topologiques.*
3. *Est-ce que G est isomorphe au produit semi-direct de $\mathbb{R}^{2n}$ par Z ?*

Transvections et applications

F.31. Exercice (Transvections, version abstraite)

Soit E un espace vectoriel de dimension $n \geqslant 2$ sur un corps $\mathbb{K}$. On rappelle qu'une transvection τ d'hyperplan H et de droite $D \subset H$ est un endomorphisme qui vérifie

$$\tau(x) = x + \ell(x)a,\ x \in E,$$

où $\ell \in E^*$ non nul et $a \in E$ non nul sont liés par la relation : $\ell(a) = 0$. Si on note $D = \langle a \rangle$ la droite engendrée par a et $H = \operatorname{Ker} \ell$, on dit que τ est une transvection d'hyperplan H et de droite $D \subset H$. Dire que τ est une transvection est équivalent à dire que τ, dans une base convenable, a une

matrice de la forme

$$\begin{pmatrix} 1 & \dots & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ \vdots & \dots & 1 & 1 \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

Le but de l'exercice est de montrer que les transvections engendrent $\mathrm{SL}(E)$.

1. *On suppose deux hyperplans distincts* H, H' *et* a *n'appartenant pas à ces deux hyperplans. On considère l'hyperplan* $(H \cap H') \oplus \mathbb{K}a$. *Montrer que l'on peut trouver* $\ell \in E^*$ *de noyau* $(H \cap H') \oplus \mathbb{K}a$ *telle qu'une transvection* τ *associée à* ℓ *comme ci-dessus envoie* H *sur* H'.

   Décomposer a dans une somme directe $(H \cap H') \oplus \mathbb{K}y \oplus \mathbb{K}y'$, où $y \in H$ et $y' \in H'$, puis imposer la condition $\tau(y) = \alpha y'$.

2. *On considère deux éléments distincts* x *et* y *non nuls de* E.

   (a) *On suppose* x *et* y *non colinéaires. Montrer qu'il existe une transvection* τ *telle que* $\tau(x) = y$.

   Prendre $a = y - x$ et H un hyperplan tel que $a \in H$ et $x \notin H$.

   (b) *On suppose* x *et* y *colinéaires. Montrer que l'on peut trouver deux transvections* τ *et* τ' *telles que* $\tau\tau'(x) = y$.

   Prendre un z non colinéaire à x et y et se reporter à question (a).

3. *Montrer par récurrence sur la dimension de* E *que* $\mathrm{SL}(E)$ *est engendré par les transvections.*

   Soit u dans $\mathrm{SL}(E)$, alors, par les questions précédentes, on peut trouver un produit de transvections σ tel que $u\sigma$ fixe un élément x de E et un hyperplan H tel que $x \notin H$.

F.32. Exercice (Connexité du groupe spécial linéaire)

En utilisant le fait que $\mathrm{SL}_n(\mathbb{R})$ *est engendré par les transvections (voir exercice précédent), montrer qu'il est connexe par arcs.*

Si un élément s'écrit, avec des notations inspirées de l'exercice F.31 : $\prod_i \tau_{\ell_i, a_i}$, on considère le chemin $t \mapsto \prod_i \tau_{t\ell_i, a_i}$.

Topologie sur les espaces quotient

F.33. Exercice (Continuité des racines par rapport aux coefficients)

Cet exercice permet de voir que l'ensemble des racines complexes avec multiplicité d'un polynôme dépend continûment des coefficients de ce polynôme.

Pour tout n-uplet $z = (z_1, \dots, z_n)$ de $\mathbb{C}^n$ et tout k de $[1, n]$, on note $e_k(z)$ la k-ième fonction symétrique élémentaire de z définie par l'égalité suivante, valable dans $\mathbb{C}[X]$:

$$(X - z_1) \times \cdots \times (X - z_n) = X^n + \sum_{k=1}^{n} (-1)^k e_k(z) X^{n-k}.$$

1. *Montrer que l'application $e : \mathbb{C}^n \to \mathbb{C}^n$, $z \mapsto \big(e_1(z), \dots, e_n(z)\big)$ est continue, surjective et que pour tout $w \in \mathbb{C}^n$, $e^{-1}(w)$ est une orbite sous le groupe symétrique $\mathfrak{S}_n$ (pour l'action par permutation).*
   Considérer le polynôme $X^n + \sum_{k=1}^{n} (-1)^k e_k(z) X^{n-k}$ et factoriser !
2. *Montrer que l'image d'un fermé F par e est un fermé.*
   Borner les racines du polynôme, puis prendre une suite (S_m) de $e(F)$ qui converge vers un w et montrer que $e^{-1}(S_m)$ possède une valeur d'adhérence.
3. *En déduire que e réalise un homéomorphisme entre $\mathbb{C}^n/\mathfrak{S}_n$ et $\mathbb{C}^n$.*
   On appliquera les résultats de la section D.1.2.

F.34. Exercice (Schémas de Hilbert)

L'objet de l'exercice est de mettre une topologie sur l'ensemble Hilb_n des idéaux I de codimension n de l'algèbre $\mathbb{C}[X, Y]$, c'est-à-dire tels que l'espace vectoriel $\mathbb{C}[X, Y]/I$ soit de dimension finie n. Pour cela, on le réalise comme un espace quotient. On considère l'ensemble

$$\mathfrak{Com} = \{(A, B) \in \mathscr{M}_n(\mathbb{C}) \times \mathscr{M}_n(\mathbb{C}),\ AB = BA\}.$$

1. *Soit P un polynôme de $\mathbb{C}[X, Y]$ et $(A, B) \in \mathfrak{Com}$. Donner un sens à $P(A, B) \in \mathscr{M}_n(\mathbb{C})$. Montrer que $\mathrm{GL}_n(\mathbb{C})$ agit diagonalement sur $\mathfrak{Com}$ par : $g \cdot (A, B) = (gAg^{-1}, gBg^{-1})$.*
2. *On dit qu'un couple (A, B) de $\mathfrak{Com}$ est* cyclique *s'il existe un vecteur (dit cyclique) v de $\mathbb{C}^n$ tel que l'on ait :*

   $$\{P(A, B)(v),\ P \in \mathbb{C}[X, Y]\} = \mathbb{C}^n.$$

   On note $\mathfrak{Com}^0$ l'ensemble des couples cycliques de $\mathfrak{Com}$. Montrer que $\mathfrak{Com}^0$ est un ouvert non vide de $\mathfrak{Com}$, stable pour l'action de GL_n.
   Pour la propriété d'être ouvert, on pourra montrer que le complémentaire de $\mathfrak{Com}^0$ est le lieu d'annulation de certains déterminants $n \times n$.
3. *Montrer que si (A, B) est cyclique, alors l'ensemble $I(A, B)$ défini par $I(A, B) = \{P,\ P(A, B) = 0\}$ est un élément de Hilb_n.*
   On voit bien que I est un idéal. Pour montrer qu'il est de codimension n, on prendra v un vecteur cyclique associé, et l'on considérera le morphisme surjectif $\mathbb{C}[X, Y] \to \mathbb{C}^n$, $P \mapsto P(A, B) \cdot v$ de noyau I.

4. *On considère donc l'application* π : $\mathfrak{Com}^0 \to \mathrm{Hilb}_n$, *qui envoie* (A, B) *sur l'idéal* $I(A, B)$. *Montrer que pour tout* I *de* $H^n(\mathbb{C})$, $\pi^{-1}(I)$ *est une* $\mathrm{GL}_n(\mathbb{C})$*-orbite.*

   On considérera $V := \mathbb{C}[X, Y]/I$, qui est isomorphe à $\mathbb{C}^n$, et μ_X, μ_Y les endomorphismes de multiplication respectivement par X et par Y. Pour (A, B) et (A', B') dans $\pi^{-1}(I)$, on pourra construire deux tels isomorphismes et, par composition, trouver un g tel que $g \cdot (A, B) = (A', B')$.

5. *Conclure : construire une topologie sur l'ensemble* Hilb_n *à l'aide de* π.

Cet exercice sera prolongé par l'exercice IX-C.3, *avec en supplément une structure de variété différentielle.*

Chapitre III

Réduction des endomorphismes

There's no place like home!

The Wizard of Oz,

Victor Fleming, 1939.

Nous allons appliquer systématiquement la méthode du premier chapitre dans le contexte de la réduction des endomorphismes : définir, en motivant son importance si nécessaire, une action continue de groupe topologique sur un espace topologique, puis trouver des invariants totaux pour cette action, et enfin étudier les adhérences d'orbites ou l'évolution de ces invariants totaux après passage à la limite.

On étudie ici l'action par conjugaison de $\mathrm{GL}_n(\mathbb{K})$ sur $\mathscr{M}_n(\mathbb{K})$. Ce problème, qui se ramène à l'étude des matrices semblables (et donc aux fameux invariants de similitude), est beaucoup plus difficile, que celui des matrices équivalentes. Tout d'abord, il va dépendre du corps choisi (mais pas dramatiquement) ; notre étude va se limiter au cas $\mathbb{K} = \mathbb{C}$. Mais surtout, la classification est radicalement différente si l'on s'intéresse aux matrices diagonalisables ou aux nilpotentes. On étudiera les deux, et le cas général se déduit alors de la décomposition de Dunford (en annexe A). L'invariant total des matrices diagonalisables est le spectre avec multiplicités, celui des matrices nilpotentes est la suite des dimensions des noyaux emboîtés. Pour l'étude topologique, il sera pratique de stocker l'information relative aux noyaux emboîtés sous forme de *diagramme de Young*.

Deux raisons précises à cela : d'une part, une vision synthétique en lignes et en colonnes ; la lecture horizontale (resp. verticale) donne des informations

sur la dimension des noyaux emboîtés (resp. la taille des blocs de Jordan) ; d'autre part, l'adhérence des orbites se décrit très facilement en termes de basculement des boîtes dans les diagrammes de Young. On peut rapprocher cette double lecture de celle d'une partition, où l'on suit la mélodie horizontalement et où l'harmonie se révèle verticalement.

1. Action $\mathrm{GL}_n(\mathbb{C})$ sur $\mathscr{D}_n(\mathbb{C})$ par conjugaison

On note $\mathscr{D}_n(\mathbb{C})$ l'ensemble des matrices diagonalisables sur $\mathbb{C}$. On rappelle que deux matrices sont semblables si et seulement si elles sont dans la même orbite de conjugaison sous l'action de $\mathrm{GL}_n(\mathbb{C})$ et que l'orbite de A sous l'action de $\mathrm{GL}_n(\mathbb{C})$ par conjugaison, aussi appelée classe de similitude de A est :

$$\mathscr{O}_A = \{PAP^{-1} \,:\, P \in \mathrm{GL}_n(\mathbb{C})\}.$$

On considère le spectre d'une matrice comme la donnée de n complexes (non nécessairement distincts) à permutation près ; un spectre sera donc vu comme un élément de l'ensemble-quotient $\mathbb{C}^n/\mathfrak{S}_n$. Il est prudent d'agiter ici une crécelle : le spectre est souvent vu comme l'ensemble des valeurs propres (sans considération de multiplicité). Dans ce chapitre, où nous nous intéresserons au problème des invariants d'actions dans un cadre algébriquement clos, il est plus judicieux d'intégrer la donnée des multiplicités dans l'ensemble $\mathbb{C}^n/\mathfrak{S}_n$.

Effectivement, si l'on considère deux matrices diagonalisables, alors elles sont semblables si et seulement si leurs valeurs propres *avec* multiplicité (on en compte alors n) sont les mêmes. Et l'on se convainc facilement que pour donner un sens et une unicité à un ensemble de valeurs propres pouvant se répéter, nous devons travailler sur $\mathbb{C}^n$, mais à permutation près. Cela va permettre dans le théorème qui suit de fournir un invariant total à l'action de $\mathrm{GL}_n(\mathbb{C})$ par conjugaison sur les matrices diagonalisables.

1.1. Théorème. *L'application*

$$\begin{aligned} \varphi : \mathscr{D}_n(\mathbb{C})/\mathrm{GL}_n(\mathbb{C}) &\longrightarrow \mathbb{C}^n/\mathfrak{S}_n \\ \mathscr{O}_A &\longmapsto \mathrm{Spec}(A) \end{aligned}$$

est bien définie et bijective.

1.2. Remarque. Le mot « réduction » prend ici tout son sens ! Puisque $\mathscr{D}_n(\mathbb{C})$ est, on l'a déjà vu et l'on le reverra, dense dans $\mathscr{M}_n(\mathbb{C})$, et contient un ouvert de $\mathscr{M}_n(\mathbb{C})$ (voir exercice D.10) ; l'espace $\mathscr{M}_n(\mathbb{C})$ est de dimension n^2, alors que $\mathbb{C}^n$ est de dimension n.

Démonstration. L'application φ est bien définie, car deux matrices semblables ont le même polynôme caractéristique et donc le même spectre. Donc, le spectre ne dépend pas du choix de l'élément A de l'orbite $\mathscr{O}_A$.

De plus, φ est bijective. En effet, soit $(\lambda_1,\dots,\lambda_n)$ un n-uplet de réels représentant un élément S de $\mathbb{C}^n/\mathfrak{S}_n$. Alors, S est l'image par φ de la matrice diagonale $D=\mathrm{diag}(\lambda_1,\dots,\lambda_n)$ dont les coefficients diagonaux sont les λ_k. De plus, si l'orbite $\mathscr{O}_A$ d'une matrice diagonalisable A admet pour image par φ la classe S, alors comme A est diagonalisable et admet pour valeur propre les λ_i, la matrice A est semblable à D, d'où $\mathscr{O}_A=\mathscr{O}_D$. L'injectivité de φ en résulte. □

Comme la donnée des valeurs propres (avec multiplicité!) à permutation près est équivalente à la donnée du polynôme caractéristique[1], nous avons l'énoncé qui suit.

1.3. Corollaire.

Le polynôme caractéristique ou le spectre, i.e. *les valeurs propres avec multiplicités, sont des invariants totaux de similitude pour les matrices diagonalisables.*

1.4. Remarque. En revanche, le polynôme minimal est un invariant mais pas un invariant total. Il est facile d'en trouver un exemple : les matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \text{ et } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

ont le même polynôme minimal mais ne sont pas semblables. On voit également l'importance de compter les multiplicités.

La question algébrique est donc plus simple que ce que l'on imaginait.

Passons à l'étude topologique : d'une part, il est naturel (les deux espaces étant munis d'une topologie) de se demander si la bijection du théorème 1.1 est un homéomorphisme. La réponse est positive!

1.5. Exercice. *En utilisant la proposition* II-A.1.2 *et l'exercice* II-F.33, *montrer que* $\mathscr{D}_n(\mathbb{C})/\mathrm{GL}_n(\mathbb{C})$ *et* $\mathbb{C}^n/\mathfrak{S}_n$ *sont homéomorphes.*

D'autre part, nous étudions ici systématiquement la décomposition des adhérences d'orbites en orbites. Et là, les choses sont encore plus simples, puisque nous allons montrer que les orbites sont fermées. Et, on obtient même une réciproque, c'est-à-dire, une caractérisation topologique de la diagonalisabilité.

[1]Notons qu'il s'agit là d'une incursion du théorème fondamental de l'arithmétique des polynômes : l'unicité de la factorisation.

1.6. Proposition

Une matrice complexe A est diagonalisable si et seulement si son orbite $\mathscr{O}_A$ sous $\mathrm{GL}_n(\mathbb{C})$ est fermée dans $\mathscr{M}_n(\mathbb{C})$.

Démonstration. Supposons $A \in \mathscr{M}_n(\mathbb{C})$ diagonalisable. Soit $(B_k)_{k\in\mathbb{N}}$ une suite de matrices de $\mathscr{O}_A$ qui converge vers B. Montrons que $B \in \mathscr{O}_A$.

Notons $\lambda_1, \ldots, \lambda_r$ les valeurs propres distinctes de A. Le polynôme minimal de A est par hypothèses scindé simple : $\mu(X) = \prod_{i=1}^{r}(X - \lambda_i)$. Pour tout k, comme B_k est semblable à A, on a : $\mu(B_k) = 0$. Par continuité de l'application polynomiale $\mathscr{M}_n(\mathbb{K}) \to \mathscr{M}_n(\mathbb{K})$, $M \mapsto \mu(M)$, on obtient à la limite : $\mu(B) = 0$. Ainsi, B, qui est annulée par un polynôme scindé à racines simples, est diagonalisable.

Pour montrer que B est dans $\mathscr{O}_A$, il suffit donc de montrer que les polynômes caractéristiques de B et de A sont égaux. Or, par hypothèse, on a pour tout k l'égalité des polynômes caractéristiques : $\chi_{B_k} = \chi_A$. Il vient par continuité de l'application « polynôme caractéristique » la relation : $\chi_B = \chi_A$ (proposition II-A.1.2), d'où l'assertion.

Inversement, supposons A non diagonalisable, et montrons que l'orbite $\mathscr{O}_A$ n'est pas fermée. Tout d'abord, nous savons que A est semblable à une matrice triangulaire inférieure $T = (t_{ij})$. Soit, pour ε non nul, la matrice de passage[2]

$$P_\varepsilon = \begin{pmatrix} 1 & & & 0 \\ 0 & \varepsilon & & \\ & & \ddots & 0 \\ 0 & & & \varepsilon^{n-1} \end{pmatrix}.$$

On a clairement $P_\varepsilon T P_\varepsilon^{-1} = (\varepsilon^{i-j} t_{ij})$. Comme par hypothèse, t_{ij} est nul dès que $i-j<0$, on voit que la matrice $\lim_{\varepsilon\to 0} PTP^{-1}$ est diagonale : elle appartient à l'adhérence de $\mathscr{O}_T = \mathscr{O}_A$, mais pas à $\mathscr{O}_A$ (car A n'est pas diagonalisable). Ainsi, $\mathscr{O}_A$ n'est pas fermée. □

1.7. Remarque. On a vu dans la proposition II-2.2.1 que $\mathscr{D}_n(\mathbb{C})$ est dense dans $\mathscr{M}_n(\mathbb{C})$; on peut montrer que son intérieur est constitué des matrices dont toutes les valeurs propres sont différentes, voir l'exercice D.10.

Sur le corps des réels, cette propriété de densité est fausse, mais l'on peut caractériser les orbites fermées : l'orbite $\mathscr{O}_A$ est fermée si et seulement si la matrice A est semi-simple : voir, dans le second tome, le §II-1 ; voir aussi l'exercice II-1.5 pour un garde-fou à toute épreuve.

[2] Cette matrice sera ton amie dans nombre d'exercices !

2. Action $\mathrm{GL}_n(\mathbb{C})$ sur $\mathcal{N}_n(\mathbb{C})$ par conjugaison

Après avoir étudié les orbites « diagonalisables », la seconde étape est naturellement l'étude de l'action du groupe $\mathrm{GL}_n(\mathbb{C})$ sur l'ensemble des matrices nilpotentes de taille n. Pour référence ultérieure, rappelons quelques définitions bien connues.

2.1. Matrices nilpotentes

2.1.1. Définition. Soit n un entier naturel non nul. Une matrice $A \in \mathcal{M}_n(\mathbb{C})$ est dite *nilpotente* s'il existe un entier naturel non nul m tel que : $A^m = 0$. On appelle *ordre de nilpotence* le plus petit entier m tel que $A^m = 0$.

On note $\mathcal{N}_n(\mathbb{C})$ l'ensemble des matrices nilpotentes de $\mathcal{M}_n(\mathbb{C})$.

L'action par conjugaison de $\mathrm{GL}_n(\mathbb{C})$ sur $\mathcal{M}_n(\mathbb{C})$ stabilise $\mathcal{N}_n(\mathbb{C})$. On appelle *orbite nilpotente* une orbite de l'action restreinte à $\mathcal{N}_n(\mathbb{C})$: plus classiquement, c'est une classe de similitude de matrices nilpotentes.

Ces définitions s'étendent à un corps quelconque.

Notons que l'ordre de nilpotence dans la définition vérifie toujours $m \leqslant n$. Effectivement, les valeurs propres de la matrice nilpotente A sont toutes nulles et donc le théorème de Cayley-Hamilton assure la relation : $A^n = 0$.

2.1.2. Mise en garde. L'ensemble $\mathcal{N}_n(\mathbb{C})$ des matrices nilpotentes n'est pas un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{C})$. Par exemple, la somme des deux matrices nilpotentes suivantes n'est pas nilpotente :

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

En revanche, $\mathcal{N}_n(\mathbb{C})$ est un cône, que l'on appellera donc cône nilpotent, c'est-à-dire qu'il est stable par homothéties : pour $A \in \mathcal{N}_n(\mathbb{C})$ et $\lambda \in \mathbb{C}$ quelconques, on a : $\lambda A \in \mathcal{N}_n(\mathbb{C})$.

Pour en savoir plus sur ce merveilleux objet de la géométrie algébrique, se référer aux paragraphes II-3 et IV-4.

2.2. Noyaux itérés et injections de Frobenius

Soit $A \in \mathcal{N}_n(\mathbb{C})$. On va noter $K_i = \mathrm{Ker}(A^i)$ ses noyaux emboîtés. On a :

$$\{0\} = K_0 \subsetneq K_1 \subset K_2 \subset \cdots \subset K_n = \mathbb{C}^n.$$

Nous allons voir que la suite des $k_i := \dim K_i$ « s'essouffle » dans la terminologie de [24].

2.2.1. Lemme (Notations ci-dessus). *On a :*

$$\forall i \in [\![1, n-1]\!], \quad 0 \leqslant \dim K_{i+1} - \dim K_i \leqslant \dim K_i - \dim K_{i-1}.$$

Démonstration. Soit i compris entre 1 et $n-1$. La première inégalité est triviale : $K_i \subset K_{i+1}$.

On a même : $AK_{i+1} \subset K_i$. On peut donc composer les morphismes :

$$\begin{array}{ccccc} K_{i+1} & \xrightarrow{\nu} & K_i & \xrightarrow{\pi_i} & K_i/K_{i+1} \\ X & \longmapsto & AX,\ Y & \longmapsto & \overline{Y}. \end{array}$$

On a : $\mathrm{Ker}(\pi_i \circ \nu) = (\pi_i \circ \nu)^{-1}(\{\bar{0}\}) = \nu^{-1}\big(\pi_i^{-1}(\{\bar{0}\})\big) = \nu^{-1}(K_{i-1}) = K_i$.
On obtient par passage au quotient une injection : $K_{i+1}/K_i \hookrightarrow K_i/K_{i-1}$.

De cette injection on déduit l'inégalité : $\dim(K_{i+1}/K_i) \leqslant \dim(K_i/K_{i-1})$, et le lemme. □

En particulier, la suite $(\dim K_i)$ est strictement croissante avant de devenir stationnaire (au pire, à partir du rang n puisque A est nilpotente). On a donc, pour un certain rang m (dont on vérifie que c'est par définition l'indice de nilpotence) :

$$\{0\} = K_0 \subsetneq K_1 \subsetneq K_2 \subsetneq \cdots \subsetneq K_m = \mathbb{C}^n.$$

Le noyau de A^i dépend bien évidemment de A, mais sa dimension ne dépend que de l'orbite de A pour la conjugaison. Posons, pour tout entier i :

$$\lambda_i = \lambda_i(A) := k_i - k_{i-1} \quad \text{où } k_i = \dim \mathrm{Ker}\, A^i.$$

2.3. Diagramme de Young associé à une orbite nilpotente

Nous avons ainsi défini, pour chaque orbite nilpotente, deux suites d'entiers positifs ou nuls à partir d'un élément quelconque A de l'orbite :
– d'une part : $k_i = \dim \mathrm{Ker}\, A^i$,
– d'autre part : $\lambda_i = k_i - k_{i-1}$.
D'après le lemme 2.2.1, on a pour tout entier i : $0 \leqslant \lambda_i \leqslant k_i$. De plus :

$$\forall i \in \mathbb{N},\ \lambda_{i+1} \leqslant \lambda_i, \quad \text{et} \quad \sum_{j=1}^{n} \lambda_j = n.$$

On reconnaît là la définition d'une partition B.1.1. On l'intègre dans l'algèbre linéaire de la façon suivante.

2.3.1. Définition. Étant donné un entier naturel non nul n et une orbite nilpotente $\mathcal{O}$ de $\mathcal{N}_n(\mathbb{C})$, on appelle *partition associée à $\mathcal{O}$ (par les noyaux itérés)* et on note $\lambda_{\mathcal{O}} = (\lambda_i)$ la partition dont les parts sont :

$$\lambda_i = \dim \mathrm{Ker}\, A^i - \dim \mathrm{Ker}\, A^{i-1} \quad (i \in \mathbb{N}^*),$$

où A est un élément quelconque de $\mathcal{O}$.

2.3.2. Définition. On appelle *diagramme de Young* associé à une orbite nilpotente $\mathcal{O}$ (par les noyaux itérés) et l'on note $Y := \mathrm{Y}(\mathcal{O})$, ou $\mathrm{Y}(A)$ pour un élément quelconque $A \in \mathcal{O}$, le diagramme de Young associé à la partition λ associée à $\mathcal{O}$ par les noyaux itérés :

$$\lambda_i = \dim \mathrm{Ker}\, A^i - \dim \mathrm{Ker}\, A^{i-1}.$$

Il est raisonnable de se demander pourquoi cette façon d'assembler des boîtes en diagrammes serait si pratique. Une première raison en est que l'on peut lire dans le diagramme une information horizontale (les dimensions des noyaux itérés) et une verticale que nous allons présenter sous peu. Une autre raison est que l'on peut mettre des informations à l'intérieur les cases, un peu comme dans un agenda.

Parmi les problèmes de conventions dans le domaine des mathématiques, il faut savoir que le diagramme de Young en possède plusieurs, dont le choix dépend du point de vue que l'on a sur l'objet. On a d'une part la convention anglaise qui consiste à justifier le diagramme en haut et à gauche, et d'autre part la convention française qui consiste à le justifier en bas et à gauche. Faute d'accords européens en la matière et en attendant le prochain sommet, nous avons choisi la seconde. Mais il y a un autre choix à faire, celui d'assigner un diagramme de Young à une orbite. Encore une fois, deux possibilités : soit celle que nous avons choisie où la suite des noyaux itérés peut s'évaluer de bas en haut, soit alors de gauche à droite. Plusieurs ouvrages choisissent cette dernière pour des raisons naturelles : on considère la suite des noyaux itérés de gauche à droite parce que l'on écrit une inclusion de gauche à droite. Mais le point de vue du présent ouvrage se focalise plus particulièrement sur les adhérences, ou dégénérescences, d'orbites ; et il nous semble plus visuellement parlant de voir (!) la dégénérescence comme une chute des boîtes du tableau, résultat qu'assure le théorème 2.7.1.

Notons enfin que l'art de remplir des diagrammes Young débouche sur la notion de tableaux de Young. Pour une introduction succincte des techniques liées à ces objets, le livre *Young Tableaux* [14] est tout indiqué.

2.4. Scindage des noyaux itérés

Une étude préliminaire s'impose avant l'énoncé du théorème de Jordan.

Soit $A \in \mathcal{M}_n(\mathbb{C})$ une matrice nilpotente d'ordre m et, comme ci-dessus, $K_i = \operatorname{Ker} A^i$ pour tout i. Ces noyaux sont emboîtés : $\{0\} \subsetneq K_1 \subsetneq \cdots \subsetneq K_m = \mathbb{C}^n$.

On part d'un sous-espace G_m supplémentaire de K_{m-1} dans $K_m = \mathbb{C}^n$, de sorte que $K_{m-1} \oplus G_m = K_m$. On pose :

$$\lambda_m := \dim G_m = \dim K_m - \dim K_{m-1}.$$

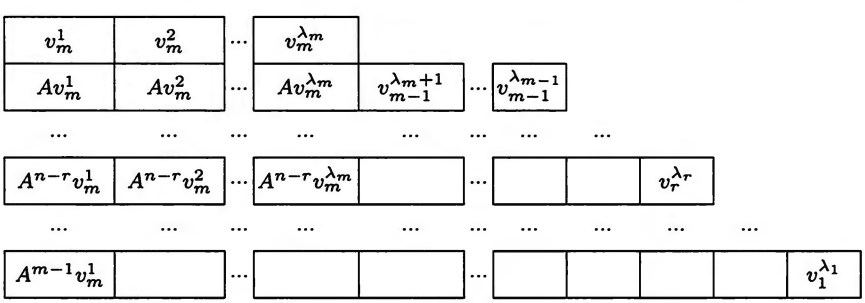
Tirons tout d'abord une remarque de la somme directe ci-dessus : Si $v \in G_m$ et $A^{m-1}v = 0$ alors $v = 0$, puisque par construction, la restriction de A^{m-1} à G_m est injective.

Soit $(v_m^1, \ldots, v_m^{\lambda_m})$ une base de G_m. Ensuite, on construit un supplémentaire G_{m-1} de K_{m-2} dans K_{m-1} de la façon suivante : on montre en utilisant la remarque précédente que $(Av_m^1, \ldots, Av_m^{\lambda_m})$ est une famille libre qui engendre un espace intersectant K_{m-2} trivialement, famille que l'on peut donc compléter en une base $(Av_m^1, \ldots, Av_m^{\lambda_m}, v_{m-1}^{\lambda_m+1} \ldots, v_{m-1}^{\lambda_{m-1}})$ d'un supplémentaire G_{m-1} de K_{m-2} dans K_{m-1}.

On note cette base $(v_{m-1}^1, \ldots, v_{m-1}^{\lambda_m}, v_{m-1}^{\lambda_m+1} \ldots, v_{m-1}^{\lambda_{m-1}})$.

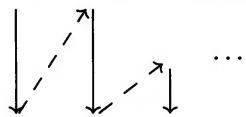
Par récurrence descendante sur r, on construit ainsi un supplémentaire G_r de K_{r-1} dans K_r de dimension $\lambda_r := \dim K_r - \dim K_{r-1}$. Par itération, on peut remplir le diagramme de Young de A ligne par ligne, successivement en multipliant par A puis en complétant.

Les amateurs de *sudoku* compléteront le diagramme de Young suivant :



Observons que l'image par A d'un vecteur est le vecteur situé dans la case en-dessous de lui, ou le vecteur nul s'il est dans la ligne du bas. On obtient

alors en lisant le tableau colonne après colonne et de haut en bas



une nouvelle base de $\mathbb{C}^n$:

$$(v_m^1, Av_m^1, \ldots, A^{m-1}v_m^1, v_m^2, \ldots, A^{m-1}v_m^2, \ldots, v_1^{\lambda_1}),$$

dans laquelle A s'écrit :

$$A' = \begin{pmatrix} J_{\lambda_1^*} & & 0 \\ & \ddots & \\ 0 & & J_{\lambda_k^*} \end{pmatrix}, \quad \text{où } J_p = \begin{pmatrix} 0 & & & 0 \\ 1 & 0 & & \\ & \ddots & 0 & \\ 0 & & 1 & 0 \end{pmatrix} \in \mathscr{M}_p(\mathbb{C}) \ (p \in \mathbb{N}^*).$$

Ici, $\lambda^* = (\lambda_j^*)_{1 \leqslant j \leqslant k}$ est la partition duale de λ (obtenue en lecture verticale, voir le paragraphe B.2). On dit alors que A' est la[3] réduite de Jordan semblable à A au sens suivant.

2.4.1. Définition. Soit p un entier naturel non nul. On appelle *bloc de Jordan nilpotent* de taille p, ou *bloc de Jordan associé à la valeur propre* 0, voir plus simplement *bloc de Jordan* si la valeur propre est sous-entendue, la matrice :

$$J_p = \begin{pmatrix} 0 & & & 0 \\ 1 & 0 & & \\ & \ddots & 0 & \\ 0 & & 1 & 0 \end{pmatrix} \in \mathscr{M}_p(\mathbb{R}).$$

On appelle *réduite de Jordan (nilpotente)* ou *forme normale de Jordan* toute matrice diagonale par blocs dont les blocs diagonaux sont des blocs de Jordan de taille décroissante.

Plus précisément, si $\nu = (\nu_i)_{i \geqslant 1}$ est une partition de n, on note J_ν la matrice diagonale dont les blocs diagonaux sont les blocs de Jordan J_{ν_i} de tailles respectives ν_1, ν_2, ...

On vient donc de démontrer la partie « existence » du théorème de Jordan. Pour être plus explicite...

[3] L'article défini est un peu hâtif pour l'instant ; on n'a pas encore montré l'unicité d'une telle matrice.

2.4.2. Proposition (Existence d'une réduite de Jordan). *Soit n un entier naturel et A une matrice nilpotente de taille n. Il existe une réduite de Jordan semblable à A.*

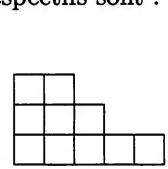
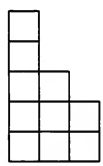
Plus précisément, soit λ la partition associée à A par les noyaux itérés, dont les parts sont :

$$\lambda_i = \dim \operatorname{Ker} A^i - \dim \operatorname{Ker} A^{i-1} \quad (i \in \mathbb{N}^*),$$

et soit $\nu = \lambda^*$. Alors, A est semblable à J_ν.

2.4.3. Exemple. Soit une matrice $A \in \mathscr{M}_{10}(\mathbb{C})$ nilpotente d'indice de nilpotence 3 telle que $\dim K_1 = 5$ et $\dim K_2 = 8$. On a donc : $\dim K_1 - \dim K_0 = 5$, $\dim K_2 - \dim K_1 = 3$, $\dim K_3 - \dim K_2 = 2$.

La partition $5 \geqslant 3 \geqslant 2$ admet pour partition duale $3 \geqslant 3 \geqslant 2 \geqslant 1 \geqslant 1$. Leurs diagrammes respectifs sont :


et

.

Donc, dans la nouvelle base, A s'écrit

$$\begin{pmatrix} J_3 & & & & \\ & J_3 & & & \\ & & J_2 & & \\ & & & J_1 & \\ & & & & J_1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & & & & & & & \\ 1 & 0 & 0 & & & & & & & \\ 0 & 1 & 0 & & & & & & & \\ & & & 0 & 0 & 0 & & & & \\ & & & 1 & 0 & 0 & & & & \\ & & & 0 & 1 & 0 & & & & \\ & & & & & & 0 & 0 & & \\ & & & & & & 1 & 0 & & \\ & & & & & & & & 0 & \\ & & & & & & & & & 0 \end{pmatrix}.$$

2.5. Classification des orbites nilpotentes

La preuve express du théorème de classification des orbites nilpotentes par les partitions montre la puissance des diagrammes de Young dans ce contexte.

2.5.1. Théorème (Jordan, formulation « classification »)

La classe de similitude d'une matrice nilpotente est caractérisée par son diagramme de Young : si A et B sont deux matrices nilpotentes de même taille, on a :

$$\mathscr{O}_A = \mathscr{O}_B \iff \mathrm{Y}(A) = \mathrm{Y}(B).$$

Démonstration. Supposons que $\mathcal{O}_A = \mathcal{O}_B$, c'est-à-dire que les matrices A et B sont semblables. Il existe alors $P \in \mathrm{GL}_n(\mathbb{C})$ telle que $B = PAP^{-1}$, d'où $\dim \mathrm{Ker}\, B^i = \dim \mathrm{Ker}(PA^iP^{-1}) = \dim \mathrm{Ker}\, A^i$ pour tout i, si bien que $\mathrm{Y}(A) = \mathrm{Y}(B)$.

Inversement, supposons que : $\mathrm{Y}(A) = \mathrm{Y}(B)$. Les matrices A et B ont alors la même partition $\lambda_A = \lambda_B$, donc la même partition duale $\lambda_A^* = \lambda_B^*$ et sont ainsi semblables à la même réduite de Jordan. Donc, A et B sont semblables, c'est-à-dire : $\mathcal{O}_A = \mathcal{O}_B$. □

Le théorème suivant est uniquement une reformulation de ce qui précède.

2.5.2. Théorème (Jordan, formulation quasi-classique)

Soit n un entier naturel non nul. Pour ν une partition de n, on note J_ν la matrice diagonale par blocs dont les blocs diagonaux sont les blocs de Jordan dont les tailles sont les parts $\nu_1, \nu_2, \ldots$ de ν.

L'application qui, à une partition $\nu \vdash n$, associe la classe de similitude $\mathcal{N}_\nu$ de J_ν, est une bijection de l'ensemble des partitions de n sur l'ensemble des classes de similitude de matrices nilpotentes de taille n.

2.5.3. Remarque (Deux conventions)

On a vu apparaître deux bijections raisonnables entre partitions et orbites nilpotentes :

- par les noyaux itérés : à $\lambda \vdash n$, on associe l'ensemble $\mathcal{O}_\lambda$ des matrices A telles que, pour tout i, on ait :

$$\dim \mathrm{Ker}\, A^i - \dim \mathrm{Ker}\, A^{i-1} = \lambda_i;$$

- par la forme normale de Jordan : à $\nu \vdash n$, on associe l'ensemble $\mathcal{N}_\nu$ des matrices semblables à J_ν.

Le lien entre ces deux bijections est la dualité des partitions. On a en effet, pour tout n :

$$\forall \lambda \vdash n,\ \mathcal{O}_\lambda = \mathcal{N}_{\lambda^*} \quad \text{et} \quad \forall \nu \vdash n,\ \mathcal{N}_\nu = \mathcal{O}_{\nu^*}.$$

Le point de vue géométrique des diagrammes de Young permet, dans une certaine mesure, de voir (au sens propre) conjointement ces deux bijections : si on lit horizontalement le tableau de Young, on obtient la partition λ ; si on le lit verticalement, on obtient la partition ν.

Encore une dernière formulation, peut-être un peu ringardisée par les précédentes, du théorème de Jordan.

2.5.4. Théorème (Jordan, formulation classique)

Soit n un entier naturel non nul et A une matrice nilpotente de taille $n\times n$. Il existe une unique suite $(\nu_1,\dots,\nu_m)$ décroissante d'entiers naturels non nuls telle que A soit semblable à la matrice J_ν.

2.5.5. Remarque. Tout ce qui précède dans le paragraphe 2, c'est-à-dire la classification des classes de similitude de matrices nilpotentes par les partitions de la taille de la matrice, reste vrai si l'on remplace $\mathbb{C}$ par un corps quelconque. Il suffit de relire les preuves en constatant que l'on n'a fait usage ni de la caractéristique zéro, ni du fait que le corps est algébriquement clos.

2.6. Aparté : dénombrement des orbites nilpotentes

Soit $p(n)$ le nombre de partitions du nombre n. Par exemple, $p(4) = 5$ correspondant aux partitions de 4 :

$$4,\quad 3\geqslant 1,\quad 2\geqslant 2,\quad 2\geqslant 1\geqslant 1,\quad 1\geqslant 1\geqslant 1\geqslant 1.$$

Ce nombre $p(n)$ compte les diagrammes de Young de taille n, et donc les classes de conjugaison de matrices nilpotentes $n\times n$. Pour le calculer, nous allons utiliser l'anneau $\mathbb{Q}[[X]]$ des séries formelles en l'indéterminée X. En effet, il est pratique de considérer la *série génératrice*

$$s(X)=\sum_{n\geqslant 0} p(n)X^n.$$

On a alors facilement ce qui suit.

2.6.1. Corollaire. *Le nombre de classes de conjugaison pour l'action de $\mathrm{GL}_n(\mathbb{C})$ sur le cône nilpotent $\mathcal{N}_n(\mathbb{C})$ est le nombre $p(n)$ égal au coefficient de X^n dans le produit infini*

$$s(X)=\prod_{k\geqslant 1}\frac{1}{1-X^k}\cdot$$

On se convaincra facilement de la formule en cherchant le terme en X^4 dans l'expression de droite que l'on calcule sous la forme

$$\prod_{k=1}^{+\infty}\frac{1}{1-X^k}=\prod_{k=1}^{+\infty}(1+X^k+(X^k)^2+\dots)=$$

$$=(1+X+X^2+\dots)\big(1+X^2+(X^2)^2+\dots\big)\big(1+X^3+\dots\big)\big(1+X^4+\dots\big)\dots$$

Le terme est donc $X^4 + XX^3 + (X^2)^2 + (X)^2X^2 + (X)^4 = 5X^4$; on voit bien comment ce produit infini « compte » les partitions de n pour tout n.

Une preuve plus complète peut être trouvée dans le livre (épuisé) d'Ivan Nourdin [28].

2.6.2. Remarque. Notons que ces séries génératrices constituent un outil précieux qui dépasse le cadre du simple rangement. Si l'on devait les vendre, le slogan en serait : « si vous pouvez pas en trouver une, vous pouvez les trouver toutes ». L'exercice classique suivant est encore un exemple édifiant.

2.6.3. Exercice. *Montrer que le nombre $I(n)$ de partitions d'un entier quelconque n dont toutes les parts sont des nombres impairs est égal au nombre $D(n)$ de partitions de n dont les parts sont des nombres distincts.*

Montrer que l'on a d'une part $\sum_{n\geqslant 0} I(n)X^n = \prod_{k\geqslant 1} 1/(1-X^{2k-1})$ et que d'autre part $\sum_{n\geqslant 0} D(n)X^n = \prod_{k\geqslant 1}(1+X^k)$. Comparer ensuite les deux produits en remarquant que tout entier s possède une unique décomposition de la forme $s = (2k-1)2^r$; on peut tout aussi bien, et d'ailleurs plus simplement, écrire : $(1+X^k) = (1-X^{2k})/(1-X^k)$.

2.7. Adhérence des orbites nilpotentes

On va déterminer l'adhérence de l'orbite $\mathcal{O}_A$ d'une matrice nilpotente complexe A. Le résultat utilise fortement l'ordre de dominance sur les diagrammes de Young introduit dans l'annexe B.3. On s'y reportera auparavant pour saisir les notations qui apparaissent dans l'énoncé ainsi que dans la démonstration qui suit.

2.7.1. Théorème. *Soit A une matrice nilpotente de taille n. Alors, l'adhérence de la classe de similitude de A est :*

$$\overline{\mathcal{O}_A} = \bigsqcup_{\mathrm{Y}(B)\geqslant \mathrm{Y}(A)} \mathcal{O}_B.$$

2.7.2. Remarque. Il en résulte, ce qui n'est pas évident *a priori* mais est arrivé souvent dans les exemples de ce livre, que l'adhérence des orbites soit caractérisée par un ordre partiel. Plus précisément, la relation suivante, définie sur l'ensemble des orbites nilpotentes, est un ordre :

$$\mathcal{O} \leqslant \mathcal{O}' \quad \text{si } \mathcal{O} \subset \overline{\mathcal{O}'}.$$

On l'appelle *ordre de Chevalley* ou *ordre de dégénérescence*. Il exprime une jolie interaction entre combinatoire et topologie.

Démonstration. On procède par double inclusion. Soit B un élément de l'adhérence de $\mathscr{O}_A$ et (A_m) une suite d'éléments de $\mathscr{O}_A$ qui converge vers B. On a pour tout i :

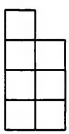
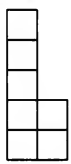
$$k_i\big(\mathrm{Y}(B)\big) = \dim \mathrm{Ker}\, B^i \geqslant \dim \mathrm{Ker}\, A_m^i = k_i\big(\mathrm{Y}(A_m)\big) = k_i\big(\mathrm{Y}(A),\big)$$

car, à la limite, le noyau ne peut pas « rétrécir » (plus formellement, c'est la semi-continuité inférieure du rang du corollaire I-4.5). Cela traduit l'inégalité de diagrammes : $\mathrm{Y}(B) \geqslant \mathrm{Y}(A)$.

On veut montrer que toute orbite de la forme $\mathscr{O}_B$ avec $\mathrm{Y}(A) \leqslant \mathrm{Y}(B)$ est incluse dans l'adhérence. Pour commencer, montrons que si $\mathrm{Y}(A) \leqslant_{el} \mathrm{Y}(B)$, alors $\mathscr{O}_B \subset \overline{\mathscr{O}_A}$.

Supposons que $\mathrm{Y}(A)$ et $\mathrm{Y}(B)$ n'ont que deux colonnes.

Soit deux entiers $p \geqslant q$, B (resp. A) une matrice dont le diagramme de Young a deux colonnes de hauteur p et q (resp. $p+1$ et $q-1$) :

B :  et A : 

Posons, pour $m \in \mathbb{N}^*$:

$$A_m = \begin{pmatrix} J_p & 0 \\ 0 & J_q \end{pmatrix} + \frac{1}{m} E_{n,p}.$$

Ici, $E_{n,p}$ est la matrice élémentaire dont le seul coefficient non nul a pour indice (n,p). Par exemple pour $n = 4+3$,

$$A_m = \begin{pmatrix} 0 & 0 & 0 & 0 & & & \\ 1 & 0 & 0 & 0 & & & \\ 0 & 1 & 0 & 0 & & & \\ 0 & 0 & 1 & 0 & & & \\ & & & & 0 & 0 & 0 \\ & & & & 1 & 0 & 0 \\ & & & \frac{1}{m} & 0 & 1 & 0 \end{pmatrix}.$$

On vérifie facilement que :

(a) $\mathrm{rg}\, A_m = n-2$;

(b) $A_m^{p+1} = 0$;

(c) $A_m^p \neq 0$.

L'assertion (a) s'établit en calculant le noyau de A_m, et l'assertion (b) en montrant que pour tout i, on a : $A^{p+1}e_i = 0$ (en fait, il suffit de le montrer

pour $i = 1$ et $i = p+1$, et le reste en découle). Pour l'assertion (c), il suffit de voir que l'on a : $A^p e_1 = \frac{1}{m} e_n \neq 0$.

On en déduit alors que $\mathrm{Y}(A_m) = \mathrm{Y}(A)$. En effet, $\mathrm{Y}(A_m)$ possède deux colonnes, celle de gauche étant de hauteur $p+1$ et la seconde s'en déduisant. Par le théorème 2.5.1, A_m est semblable à A pour tout m et donc : $\mathcal{O}_A = \mathcal{O}_{A_m}$. Comme on a :

$$\lim_{m \to +\infty} A_m = \begin{pmatrix} J_p & 0 \\ 0 & J_q \end{pmatrix} = B' \sim B,$$

on a bien : $B' \in \overline{\mathcal{O}_{A'}} = \overline{\mathcal{O}_A}$. Or, $\overline{\mathcal{O}_A}$ est stable par l'action de $\mathrm{GL}_n(\mathbb{C})$, par continuité de l'action, il vient : $\mathcal{O}_B \subset \overline{\mathcal{O}_A}$.

Regardons à présent le cas où $\mathrm{Y}(A)$ et $\mathrm{Y}(B)$ ont plus de deux colonnes. Comme $\mathrm{Y}(A) \leqslant_{el} \mathrm{Y}(B)$, toutes les colonnes sauf deux restent inchangées, il suffit donc de construire la même suite (A_m) en rajoutant éventuellement des blocs de Jordan constants (par rapport à m) lorsque la colonne correspondante est inchangée.

En conclusion, si $\mathrm{Y}(A) \leqslant_{el} \mathrm{Y}(B)$, alors $\mathcal{O}_B \subset \overline{\mathcal{O}_A}$.

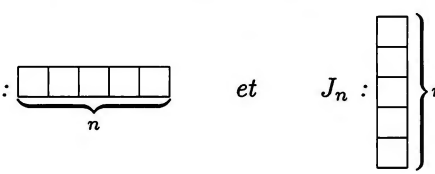
Mais alors, si $\mathrm{Y}(A) \leqslant \mathrm{Y}(B)$, on peut trouver des diagrammes de Young $Y_0, \ldots, Y_r$ et des matrices $B_0, \ldots, B_r$ telles que

$$\mathrm{Y}(A) = Y_r = \mathrm{Y}(B_r) \leqslant_{el} Y_{r-1} = \mathrm{Y}(B_{r-1}) \leqslant_{el} \cdots \leqslant_{el} Y_0 = \mathrm{Y}(B_0) = \mathrm{Y}(B).$$

On a alors : $\mathcal{O}_B = \mathcal{O}_{B_0} \subset \overline{\mathcal{O}_{B_1}} \subset \overline{\overline{\mathcal{O}_{B_2}}} \subset \cdots \subset \overline{\overline{\vdots}\,\mathcal{O}_{B_r}} = \overline{\mathcal{O}_A}$, ce qui termine la preuve du théorème. □

On déduit du théorème suivant.

2.7.3. Corollaire. *L'orbite nulle est la seule orbite fermée. L'orbite du bloc de Jordan de taille maximale est la seule orbite ouverte. Elle sont caractérisées par les diagrammes :*



Que dire maintenant des propriétés topologiques des autres orbites ? Le mieux que l'on puisse dire simplement en général, c'est qu'elles sont toujours localement fermées.

2.7.4. Corollaire. *Toute classe $\mathscr{O}$ de similitude nilpotente est localement fermée, c'est-à-dire que $\mathscr{O}$ est ouverte dans $\overline{\mathscr{O}}$.*

Démonstration. Soit Y le diagramme de Young associé à l'orbite nilpotente $\mathscr{O} = \mathscr{O}_Y$ via le théorème 2.5.1. On veut montrer que $\mathscr{O}$ est ouverte dans $\overline{\mathscr{O}}$, c'est à dire que $\overline{\mathscr{O}} \setminus \mathscr{O}$ est fermé dans $\overline{\mathscr{O}}$. Or, d'après le théorème qui précède,

$$\overline{\mathscr{O}} \setminus \mathscr{O} = \bigsqcup_{Y' > Y} \mathscr{O}_{Y'},$$

qui est fermé dans $\mathscr{M}_n(\mathbb{C})$ et donc dans $\overline{\mathscr{O}}$ par transitivité de l'ordre sur les diagrammes. □

3. Action $\mathrm{GL}_n(\mathbb{C})$ sur $\mathscr{M}_n(\mathbb{C})$ par conjugaison

Revenons au cas général, c'est-à-dire l'action de $\mathrm{GL}_n(\mathbb{C})$ sur $\mathscr{M}_n(\mathbb{C})$ par conjugaison. L'idée est de se servir des classifications du cas diagonalisable (par le polynôme caractéristique) et du cas nilpotent (diagrammes de Young) et de conclure par la décomposition de Dunford. Il y a tout de même un piège...

3.1. Remarque. Nous verrons dans le théorème qui suit que lorsque l'on écrit les décompositions de Dunford de deux matrices, $A = D + N$ et $A' = D' + N'$, alors, le fait que A est semblable à A' implique que D est semblable à D' et N à N'.

Mais la réciproque est fausse comme le prouve le contre-exemple suivant.

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ et } \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ ne sont pas semblables,}$$

puisqu'elles n'ont pas le même polynôme minimal.

Il faut donc avoir des hypothèses un peu plus restrictives, non pas sur la décomposition de Dunford dans sa globalité, mais localement, c'est-à-dire pour chaque sous-espace caractéristique. On notera dans le théorème qui suit D_λ et N_λ les restrictions respectives de D et N au sous-espace caractéristique associé à la valeur propre λ de A.

3.2. Théorème. *Soit $A = D + N$ et $A' = D' + N'$ les décompositions de Dunford de deux matrices complexes. Si A est semblable à A', alors D est semblable à D' et N est semblable à N'.*

Réciproquement, si D est semblable à D' et si, pour toute valeur propre λ de A (donc de A'), on a : $\mathrm{Y}(N_\lambda) = \mathrm{Y}(N'_\lambda)$, alors A est semblable à A'.

Démonstration. Supposons que A et A' sont semblables et soit P inversible telle que $A' = PAP^{-1}$. Il vient $D' + N' = PDP^{-1} + PNP^{-1}$. On vérifie que le membre de droite est également une décomposition de Dunford (une diagonalisable plus une nilpotente qui commutent). L'unicité de la décomposition de Dunford implique le résultat.

Passons à la réciproque. Comme D et D' sont semblables, elles ont même spectre, même polynôme caractéristique (il en est donc de même pour A et A') et de plus, pour toute valeur propre λ, les sous-espaces caractéristiques E_λ et E'_λ sont de même dimension, c'est-à-dire la multiplicité de λ dans $\chi_A = \chi_{A'}$. Il existe donc pour tout λ un isomorphisme $\varphi_\lambda : E_\lambda \to E'_\lambda$. Par hypothèse sur les diagrammes de Young, $\varphi_\lambda N_\lambda \varphi_\lambda^{-1}$ et N'_λ sont semblables, et l'on peut donc supposer que l'isomorphisme φ_λ vérifie $\varphi_\lambda N_\lambda \varphi_\lambda^{-1} = N'_\lambda$. De plus, comme D_λ et D'_λ sont des homothéties, il vient $\varphi_\lambda(D_\lambda + N_\lambda)\varphi_\lambda^{-1} = D'_\lambda + N'_\lambda$. Et ce pour tout λ, ce qui implique que A et A' sont semblables. □

Reste à comprendre la topologie des orbites. Pour cela, on peut fixer un polynôme $\chi = \prod(X - \lambda_i)^{n_i}$ de degré n et se restreindre à l'action de $\mathrm{GL}_n(\mathbb{C})$ sur

$$\mathscr{M}_n(\chi) = \{B \in \mathscr{M}_n(\mathbb{C}) \,:\, \chi_B = \chi\},$$

l'ensemble des matrices qui ont pour polynôme caractéristique χ. On note que $\mathscr{M}_n(\chi)$ est un fermé de l'espace des matrices, puisque, comme on l'a déjà vu (II-A.1.2), l'application $A \mapsto \chi_A$ est continue. Le théorème suivant en résulte.

3.3. Théorème. *Avec les notations précédentes, A appartient à l'adhérence de l'orbite de A' si et seulement si $\chi_A = \chi_{A'}$ et, pour tout élément λ du spectre, on a : $\mathrm{Y}(N_\lambda) \geqslant \mathrm{Y}(N'_\lambda)$.*

Démonstration. Laissée en exercice. □

3.4. Exercice. *Combien y a-t-il d'orbites pour l'action de $\mathrm{GL}_n(\mathbb{C})$ sur $\mathscr{M}_n(\chi)$? Retrouver à l'aide du théorème qu'une orbite $\mathscr{O}_A$ est fermée si et seulement si A est diagonalisable. Décrire l'adhérence de l'orbite $\mathscr{O}_A$, où A est une matrice quelconque. On se ramènera à l'aide du lemme des noyaux au cas où le spectre de A est réduit à un seul élément.*

4. Passage de $\mathbb{C}$ à $\mathbb{R}$

Nous venons de voir comment la classification des orbites pour l'action de similitude pouvait se résoudre sur le corps des complexes. Il est naturel de chercher maintenant à résoudre le problème analogue sur le corps des réels. Le résultat est finalement très simple, puisqu'un résultat classique dit que deux matrices réelles sont semblables sur $\mathbb{C}$ si et seulement si elles sont semblables sur $\mathbb{R}$. Il est utile de bien connaître cette preuve qui est une belle application de la polynomialité du déterminant.

4.1. Proposition. *Pour des matrices réelles,*

$$\mathrm{GL}_n(\mathbb{C})-\text{semblables} \iff \mathrm{GL}_n(\mathbb{R})-\text{semblables}.$$

Démonstration. La réciproque étant claire, il suffit de montrer que $\mathrm{GL}_n(\mathbb{C})$-semblable implique $\mathrm{GL}_n(\mathbb{R})$-semblable. On suppose donc A et B deux matrices réelles telles que $PAP^{-1} = B$ pour une matrice P de $\mathrm{GL}_n(\mathbb{C})$. Posons donc $P = U + iV$ avec U et V réelles, il vient donc $(U+iV)A = B(U+iV)$ et en prenant la partie réelle et la partie imaginaire, on obtient deux égalités $UA = BU$ et $VA = BV$. Malheureusement, rien ne garantit que U ou V soient inversibles. Il faut donc travailler un peu plus sur ce résultat. Soit t quelconque dans $\mathbb{R}$, on a donc $(U + tV)A = B(U + tV)$. Montrons qu'il existe t tel que $U + tV$ est inversible. Cela établira notre proposition puisqu'en posant $Q := U + tV$, on aura bien Q réelle et $QAQ^{-1} = B$.

Soit $F(z) = \det(U + zV)$, $z \in \mathbb{C}$. Par polynomialité du déterminant, $F(z)$ est une fonction polynomiale en z. Or $F(i) \neq 0$, donc le polynôme F est non nul, il possède donc un nombre fini de zéros, et donc il existe un réel t tel que $F(t) \neq 0$. Dit autrement, $Q = U + tV$ est inversible. □

4.2. Corollaire. *L'orbite* $\mathrm{GL}_n(\mathbb{R}) \cdot A$ *d'une matrice réelle* A *est exactement l'intersection de* $\mathrm{GL}_n(\mathbb{C}) \cdot A$ *avec* $\mathscr{M}_n(\mathbb{R})$ *:*

$$\mathrm{GL}_n(\mathbb{R}) \cdot A = \mathrm{GL}_n(\mathbb{C}) \cdot A \cap \mathscr{M}_n(\mathbb{R}).$$

Ce sera loin d'être le cas dans le chapitre V consacré à l'action de congruence, où l'intersection de $\mathrm{GL}_n(\mathbb{C}) \cdot A$ avec $\mathscr{M}_n(\mathbb{R})$ sera une réunion disjointe d'orbites pour l'action de $\mathrm{GL}_n(\mathbb{R})$.

Ces résultats seront généralisés pour tout corps dans le second tome, § II-1 sur la semi-simplicité. Nous verrons alors que la semi-simplicité généralise la notion de diagonalisabilité lorsque le corps n'est pas algébriquement clos. En particulier, l'utilisation de la décomposition de Dunford, qui a fait ses preuves, se fait dans ce cadre.

5. Invariants de similitude sur un corps quelconque

Cette partie de la théorie est un peu plus ardue et souvent mal assimilée par les étudiants. Elle est pourtant indispensable si l'on veut traiter le problème de classification qui nous intéresse dans le cadre d'un corps quelconque et si l'on veut comprendre en profondeur les problèmes de commutants. Plutôt que de regarder les invariants de similitude par la théorie des modules sur $\mathbb{K}[X]$ (même si la théorie est très implicite dans les preuves), nous choisissons de les introduire par le biais d'une méthode *light* : la décomposition de Frobenius. Nous ferons ensuite le lien entre les invariants ainsi obtenus et les diagrammes de Young.

Dans toute la suite, on travaillera dans un espace vectoriel de dimension finie n sur un corps $\mathbb{K}$ muni d'un endomorphisme u. Nous commençons par construire un raffinement du polynôme minimal μ_u de u, élément essentiel de la décomposition qui va suivre.

5.1. Lemme. *Pour tout x de E, il existe un unique polynôme unitaire $\mu_{u,x}$ tel que*

$$P(u)(x) = 0 \Leftrightarrow \mu_{u,x} | P.$$

En particulier, on a : $\mu_{u,x} | \mu_u$. Le degré de $\mu_{u,x}$ est la dimension du sous-espace $\mathbb{K}[u] \cdot x$ engendré par la famille $\big(x, u(x), \ldots, u^k(x), \ldots\big)$

Démonstration. Soit x un vecteur de E. L'ensemble des P de $\mathbb{K}[X]$ tels que $P(u)(x) = 0$ est un idéal de $\mathbb{K}[X]$. Cet idéal est principal ($\mathbb{K}$ est un corps !) et non nul puisqu'il contient μ_u. D'où l'existence et l'unicité d'un polynôme unitaire[4] $\mu_{u,x}$. Par définition, le degré de $\mu_{u,x}$ est égal au plus petit entier k tel que $\big(x, u(x), \ldots, u^k(x)\big)$ soit liée ou, dit autrement, le plus grand entier k tel que $\big(x, u(x), \ldots, u^{k-1}(x)\big)$ soit libre. Par une récurrence classique, on obtient que pour tout entier m supérieur ou égal à k, le vecteur $u^m(x)$ est une combinaison linéaire de cette famille et donc, la famille $\big(x, u(x), \ldots, u^{k-1}(x)\big)$ est une base de $\mathbb{K}[u] \cdot x$. □

5.2. Remarque. Cette dernière assertion montre en particulier que le polynôme $\mu_{u,x}$ ne dépend pas du corps de base dans le sens suivant : si u est un endomorphisme de $\mathbb{K}^n$ et si $\mathbb{L}$ est une extension de $\mathbb{K}$, alors pour tout x de $\mathbb{K}^n$ on a $\mu_{u,x} = \mu_{u_{\mathbb{L}},x}$, où $u_{\mathbb{L}}$ est l'endomorphisme étendu naturellement sur $\mathbb{L}^n$ (par la base canonique de $\mathbb{K}^n$, on peut identifier $u \mapsto u_{\mathbb{L}}$ avec l'injection naturelle $\mathscr{M}_n(\mathbb{K}) \hookrightarrow \mathscr{M}_n(\mathbb{L})$). En effet, l'invariance du rang par

[4] Il n'est peut-être pas mauvais de rappeler que si a et a' engendrent chacun le même idéal (principal) I de l'anneau A, alors $a' = \varepsilon a$ avec ε inversible dans A.

changement de corps implique que la dimension sur $\mathbb{K}$ de $\mathbb{K}[u]\cdot x$ est égale à dimension sur $\mathbb{L}$ de $\mathbb{L}[u]\cdot x$. Ainsi, les degrés des polynômes $\mu_{u,x}$ et $\mu_{u_\mathbb{L},x}$ sont égaux, et comme $\mu_{u_\mathbb{L},x}$ divise $\mu_{u,x}$ par construction, on a l'égalité.

De même, le polynôme minimal ne dépend pas du corps de base (comme le polynôme caractéristique d'ailleurs). C'est un exercice classique sur l'invariance du rang d'un système. Pour en donner un argument rapide, disons que d'une part $\mu_{u_\mathbb{L}}$ divise μ_u puisque $\mu_u(u_\mathbb{L}) = \mu_u(u)_\mathbb{L} = 0$, et d'autre part le degré $d_\mathbb{L}$ de $\mu_{u_\mathbb{L}}$ ne pourrait être strictement inférieur au degré d de μ_u, sinon l'on aurait une dépendance sur $\mathbb{L}$ entre entre les $d_\mathbb{L}$ premières puissances u^k, donc ces morphismes seraient dépendants sur $\mathbb{K}$.

5.3. Définition. Les sous-espaces $\mathbb{K}[u]\cdot x$, pour $x \in E$, sont appelés sous-espaces u-cycliques ou tout simplement sous-espaces *cycliques*. Autrement dit, un sous-espace F est cyclique s'il existe x dans E tel que $F = \mathbb{K}[u]\cdot x$. Le sous-espace F est alors le plus petit espace stable par u et contenant x.

Le polynôme minimal de u peut se réaliser comme un $\mu_{u,x}$.

5.4. Proposition. *Il existe x dans E tel que $\mu_u = \mu_{u,x}$.*

Démonstration. Soit $\mu_u = \prod_i P_i^{\alpha_i}$ la décomposition de μ_u en irréductibles. On pose $K_i := \operatorname{Ker} P_i^{\alpha_i}(u)$, qui est stable par u. Par le lemme des noyaux on a $E = \oplus_i K_i$. On pose alors $u_i := u\,|_{K_i}$ qui est un endomorphisme de K_i vérifiant $\mu_{u_i} = P_i^{\alpha_i}$.

Montrons tout d'abord le résultat de façon locale, c'est-à-dire pour u_i sur K_i ; on veut donc montrer qu'il existe x_i tel que $\mu_{u_i,x_i} = \mu_{u_i}$. Par l'absurde, pour tout x_i de K_i, μ_{u_i,x_i} diviserait strictement $\mu_{u_i} = P_i^{\alpha_i}$ donc diviserait $P_i^{\alpha_i-1}$, par irréductibilité de P_i. Mais, dans ce cas, $P_i^{\alpha_i-1}(u_i)$ serait nul sur tout K_i ce qui est impossible par minimalité de μ_{u_i}. Donc, pour tout i il existe x_i dans K_i tel que $\mu_{u_i,x_i} = \mu_{u_i}$.

Comme d'habitude, le lemme des noyaux nous permet un passage du local au global : soit $x = \sum_i x_i$. Montrons que $\mu_u = \mu_{u,x}$. Il suffit de montrer que $\mu_u | \mu_{u,x}$ puisque l'inégalité inverse (pour la divisibilité !) est assurée. Or, $0 = \mu_{u,x}(u)(x) = \mu_{u,x}(u)(\sum_i x_i)$, ce qui donne $\mu_{u,x}(u)(x_i) = 0$ pour tout i (les K_i sont en somme directe) et donc $P_i^{\alpha_i} = \mu_{u_i,x_i} | \mu_{u,x}$ pour tout i. Ainsi, comme les μ_{u_i,x_i} sont premiers entre eux, leur produit μ_u divise $\mu_{u,x}$. □

Cette proposition nous permet d'obtenir une caractérisation importante des espaces u-cycliques. On rappelle qu'une matrice compagnon associée au polynôme unitaire $P = X^n - a_{n-1}X^{n-1} - \cdots - a_0 \in \mathbb{K}[X]$ est la matrice

de $\mathscr{M}_n(\mathbb{K})$ donnée par

$$C_P = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & a_0 \\ 1 & 0 & \cdots & 0 & 0 & a_1 \\ 0 & 1 & \ddots & 0 & 0 & a_2 \\ \vdots & 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & 1 & 0 & a_{n-2} \\ 0 & 0 & \cdots & 0 & 1 & a_{n-1} \end{pmatrix}.$$

On montre que son polynôme caractéristique est P (voici, une fois n'est pas coutume, un petit conseil taupinal : si l'on ne veut pas faire de récurrence laborieuse, il faut développer par rapport à la dernière colonne), de sorte que l'on peut voir la matrice compagnon comme un choix dans l'image réciproque du singleton $\{P\}$ par l'application $A \mapsto \chi_A$.

5.5. Proposition. *Les conditions suivantes sont équivalentes :*

(i) *l'espace E lui-même est u-cyclique ;*

(ii) *il existe une base de E telle que la matrice de u soit une matrice compagnon ;*

(iii) $\mu_u = \chi_u$.

Démonstration. (i) $\Rightarrow$ (ii). Si E est u-cyclique, de dimension n, alors il existe x tel que $E = \mathbb{K}[u] \cdot x$. La famille $\mathscr{B}_x = \big(x, u(x), \ldots, u^{n-1}(x)\big)$ est génératrice : utiliser Cayley-Hamilton pour montrer que $u^n(x)$ s'écrit comme combinaison des vecteurs de $\mathscr{B}_x$, puis une récurrence pour montrer qu'il en est de même pour $u^m(x)$, avec $m \geqslant n$. Donc, $\mathscr{B}_x$ est une base par un argument de cardinalité. On obtient facilement que la matrice de u dans la base $\mathscr{B}_x$ est la matrice compagnon C_{μ_u}.

(ii) $\Rightarrow$ (iii). Comme on l'a indiqué plus haut, le polynôme caractéristique de C_P est P. Montrons que c'est aussi son polynôme minimal. Par Cayley-Hamilton, il suffit de montrer que le degré du polynôme minimal de C_P est égal au degré de P. Par l'absurde, sinon, il lui serait strictement inférieur, et on aurait ainsi un polynôme $Q = X^k - b_{k-1}X^{k-1} - \cdots - b_0$, $k < n$, qui annulerait C_P. On aurait ainsi dans la base $(e_i)_{1 \leqslant i \leqslant n}$ de l'hypothèse (ii) : $e_{k+1} = b_{k-1}e_k + \cdots + b_0 e_1$, ce qui donnerait une relation de dépendance linéaire absurde.

(iii) $\Rightarrow$ (i). C'est un corollaire immédiat de la proposition précédente puisque, dans ce cas, le degré de μ_u vaut n. □

Pour tout x de E, le sous-espace $\mathbb{K}[u] \cdot x$ est stable par u. Toutefois, il n'est pas clair que $\mathbb{K}[u] \cdot x$ possède un supplémentaire stable par u (et c'est

en général faux)[5]. Toutefois, cela devient vrai si x est un élément comme celui que nous venons de construire ci-dessus, dont le polynôme $\mu_{u,x}$ est maximal. Ceci va être fondamental pour la décomposition de E en sous-espaces u-cycliques. On pourra voir dans cette preuve une belle application de la dualité.

5.6. Proposition. *Soit x tel que $\mu_{u,x} = \mu_u$. Alors, $\mathbb{K}[u] \cdot x$ possède un supplémentaire u-stable.*

Démonstration. On appellera d le degré de μ_u. On a donc, d'après ce qui précède, une base de $\mathbb{K}[u] \cdot x$ donnée par $\mathscr{B}_x = \big(x, u(x), \dots, u^{d-1}(x)\big)$. Soit $\varphi \in E^*$ tel que

$$\varphi(x) = \varphi\big(u(x)\big) = \cdots = \varphi\big(u^{d-2}(x)\big) = 0 \ \text{ et } \ \varphi\big(u^{d-1}(x)\big) = 1.$$

La famille $(\varphi, \varphi \circ u, \dots, \varphi \circ u^{d-1})$ est une famille libre de E^* ; il suffit pour voir cela d'appliquer l'égalité $\sum_{k=0}^{d-1} a_k \varphi \circ u^k = 0$ à $u^i(x)$ pour i de 0 à $d-1$: on obtient alors successivement $a_{d-1} = 0$, $a_{d-2} = 0, \cdots, a_0 = 0$. Soit donc Φ le sous-espace vectoriel de E^* engendré par cette famille et considérons $F = \Phi^\perp \subset E$, qui est donc de dimension $n - d$.

Montrons que F est un supplémentaire u-stable de $\mathbb{K}[u] \cdot x$.

• F est u-stable.

Soit $y \in F$, alors par construction, $\varphi\big(u^k\big(u(y)\big)\big) = 0$ pour k allant de 0 à $d-2$. De plus, $\varphi\big(u^{d-1}\big(u(y)\big)\big) = \varphi\big(u^d(y)\big) = 0$ puisque, le polynôme minimal de u étant de degré d, $u^d(y)$ s'écrit en fonction des $u^k(y)$, $k < d$. Conclusion, $u(y) \in \Phi^\perp = F$.

• $F \cap \mathbb{K}[u] \cdot x = \{0\}$.

On suppose donc $y = \sum_{k=0}^{d-1} a_k u^k(x) \in F$. En appliquant $\varphi \circ u^i$ pour i allant de 0 à $d-1$, on obtient que les a_k sont tous nuls.

On a montré que $\dim F = n - \dim(\mathbb{K}[u] \cdot x)$, ce qui achève la preuve. □

Voici maintenant le théorème fondamental qui fournit les *invariants de similitude* (appelés aussi *facteurs invariants*).

[5] On peut voir les choses de cette manière : dans les axiomes qui définissent un espace vectoriel, on peut sans mal remplacer le corps de base par un anneau unitaire A et l'on définit ainsi un A-module. Ici, E est muni d'une action de $A = \mathbb{K}[u]$ (ou même d'une action de $\mathbb{K}[X]$ via l'évaluation en u). Ce langage est pratique car sous-A-module signifie alors sous-espace u-stable. Le drame fondamental des A-modules est que le théorème de la base incomplète n'y est plus vrai et que donc, dans ce contexte, le supplémentaire se fait plutôt rare.

5.7. Théorème. *Il existe $s > 0$ et une unique famille de polynômes unitaires non constants $F_1, \dots, F_s$ de $\mathbb{K}[X]$ tels que l'on ait la décomposition suivante dite de Frobenius :*

1. $F_s | F_{s-1} | \cdots | F_1$,
2. $E = \oplus_{k=1}^{s} \mathbb{K}[u] \cdot x_k$, *avec* $\mu_{u,x_k} = F_k$.

Démonstration. L'existence découle de ce qui précède : on choisit pour vecteur x_1 un vecteur x tel que $\mu_{u,x} = \mu_u$, et l'on pose $F_1 = \mu_u$, de sorte que $E = \mathbb{K}[u] \cdot x \oplus F$ avec F stable par u, d'après la proposition qui précède. De plus le polynôme minimal de $u|_F$ divise μ_u puisque μ_u annule $u|_F$. On finit facilement l'assertion par une récurrence, qui nous donne une décomposition de Frobenius de E.

Afin de montrer l'unicité des polynômes F_k (et donc de s qui sera le nombre maximal tel que F_k soit non trivial), nous allons les caractériser. Pour cela nous introduisons l'espace $\mathrm{Alt}_u^k(E,E)$ des applications k-linéaires alternées, u-invariantes de E^k dans E. C'est-à-dire l'espace des applications φ de E^k dans E telles que

1. φ est linéaire sur chaque composante de E,
2. $\varphi(y_1, \cdots, y_k) = 0$ dès que $y_i = y_j$ pour $i \neq j$,
3. $\varphi(y_1, \cdots, y_{i-1}, u(y_i), y_{i+1}, \cdots, y_k) = u\big(\varphi(y_1, \cdots, y_{i-1}, y_i, y_{i+1}, \cdots, y_k)\big)$, pour tout i et tout k-uplet $(y_1, \cdots, y_k)$ de E^k.

On vérifie qu'il s'agit bien d'un espace vectoriel. De plus, on définit pour tout polynôme P de $\mathbb{K}[X]$ et tout φ de $\mathrm{Alt}_u^k(E,E)$, l'application $P \cdot \varphi$ donnée par

$$P \cdot \varphi(y_1, \cdots, y_k) = P(u)\big(\varphi(y_1, \cdots, y_k)\big).$$

On vérifie facilement que $P \cdot \varphi$ définit encore une application k-linéaire alternée, u-invariante de E^k dans E et que, φ étant fixé, l'ensemble des polynômes P tels que $P \cdot \varphi = 0$ est un idéal de $\mathbb{K}[X]$ que l'on note $\mathscr{I}_\varphi$.

Le théorème résulte alors de la caractérisation suivante de F_k, valable pour tout k. □

5.8. Lemme. *Le polynôme F_k est l'unique générateur unitaire de l'idéal $\bigcap_{\varphi \in \mathrm{Alt}_u^k(E,E)} \mathscr{I}_\varphi$ de $\mathbb{K}[X]$.*

Démonstration. Il suffit de montrer que F_k engendre cet idéal et l'unicité sera automatiquement assurée, car F_k est unitaire. Pour cela, nous montrons dans un premier temps que F_k est dans cet idéal, puis, dans un deuxième temps, que F_k divise tout polynôme de l'idéal. Nous fixons dans la suite une décomposition de Frobenius $E = \oplus_{i=1}^{s} \mathbb{K}[u] \cdot x_i$.

Montrons que pour tout φ dans $\mathrm{Alt}_u^k(E,E)$, $F_k \cdot \varphi = 0$. Comme $F_k.\varphi$ est u-invariante, il suffit de montrer que $F_k \cdot \varphi(x_{i_1}, \cdots, x_{i_k})$ s'annule pour tout k-uplet $(i_1, \cdots, i_k)$ d'éléments de $[1,s]$ et comme φ est alternée, on peut les supposer tels que $1 \leqslant i_1 < \cdots < i_k \leqslant s$. Mais cette annulation est claire par

$$F_k \cdot \varphi(x_{i_1}, \cdots, x_{i_k}) = F_k(u)\big(\varphi(x_{i_1}, \cdots, x_{i_k})\big) = \varphi(x_{i_1}, \cdots, F_k(u)(x_{i_k})),$$

qui est nul puisque $F_k(u)(x_{i_k}) = 0$ (vu que $i_k \geqslant k$) et que φ est linéaire en la dernière composante.

Soit maintenant P annulant tout élément de $\mathrm{Alt}_u^k(E,E)$. Nous voulons montrer que F_k divise P. On va construire tout d'abord une application particulière φ_0 de $\mathrm{Alt}_u^k(E,E)$. Pour cela, on note que la décomposition de Frobenius fournit une application $\mathbb{K}$-linéaire de E dans $\oplus_{i=1}^s \mathbb{K}[X]/(F_i)$, qui à y dans E associe le s-uplet $\big(\overline{P_i(y)}\big) 1 \leqslant i \leqslant s$, où $\overline{P_i(y)}$ est la classe modulo (F_i) de polynômes $P_i(y)$ tels que

$$y = \sum_{i=1}^{s} P_i(y)(u)(x_i)$$

est la décomposition de Frobenius de y.

On va alors considérer le k-uplet $C_k(y) := \big(\overline{P_i(y)}\big)_{1 \leqslant i \leqslant k} \in \mathbb{K}[X]/(F_k)$, où l'on considère les classes des $P_i(y)$ modulo F_k, ce qui est possible car $i \leqslant k$ implique F_k divise F_i. On peut associer au k-uplet $(y_1, \cdots, y_k)$ de E_k le déterminant $D_0(y_1, \cdots, y_k)$ de la matrice dont les colonnes sont les $C_k(y_i)$, pour i de 1 à k. Il s'agit du déterminant d'une matrice à coefficient dans l'anneau $\mathbb{K}[X]/(F_k)$, et l'on définit donc bien une application de E^k dans $\mathbb{K}[X]/(F_k)$ qui à $(y_1, \cdots, y_k)$ associe $D_0(y_1, \cdots, y_k)$. Comme x_k est annulé par $F_k(u)$, il fait sens de définir $\varphi_0(y_1, \cdots, y_k) = D_0(y_1, \cdots, y_k)(u)(x_k)$. Les propriétés classiques du déterminant sur un anneau (forme alternée, multilinéarité sur l'anneau) donnent facilement que φ_0 est bien dans $\mathrm{Alt}_u^k(E,E)$. Et donc, P annule φ_0.

De plus, on trouve (en calculant le déterminant de la matrice identité !) $\varphi_0(x_1, \cdots, x_k) = x_k$. Donc, $P(u)$ annule x_k et donc F_k divise P comme désiré, par construction de F_k. Ce qui achève la preuve du lemme et, vraiment, du théorème. □

Ce serait dommage de laisser un beau théorème comme cela sans corollaire. En voici pour commencer une formulation très pratique.

5.9. Corollaire. *Soit u un endomorphisme de E. Alors, il existe une unique famille de polynômes unitaires non constants $F_1, \ldots, F_s$ de $\mathbb{K}[X]$ tels que :*

1. *$F_s | F_{s-1} \ldots | F_1$;*
2. *il existe une base de E telle que la matrice de u dans cette base soit la matrice diagonale par blocs $(C_{F_1}, C_{F_2}, \ldots, C_{F_s})$.*

Il est clair par ce qui précède que les facteurs invariants constituent un invariant total de similitude (c'est quand même un minimum). En effet, l'unicité permet de montrer que c'est invariant par changement de base et la forme compagnon par blocs montre que l'invariant est total. On pourra ainsi parler d'invariant de similitude d'une matrice.

5.10. Corollaire. *Deux matrices sont semblables si et seulement si elles ont les mêmes facteurs invariants.*

On a vu à la remarque 5.2 que le polynôme minimal ne dépendait pas du corps de base. En fait, il en découle par construction que les invariants de similitude ne changent pas non plus lorsque l'on travaille sur une extension.

5.11. Corollaire. *Soit u un endomorphisme de $\mathbb{K}^n$ et soit $\mathbb{L}$ une extension de $\mathbb{L}$. Alors, u et $u_{\mathbb{L}}$ ont les mêmes facteurs invariants.*

Démonstration. Appelons $F_1, \ldots, F_s \in \mathbb{K}[X]$ les invariants de similitude de u et $F_1^{\mathbb{L}}, \ldots F_{s'}^{\mathbb{L}} \in \mathbb{L}[X]$ les invariants de similitude de $u_{\mathbb{L}}$. La matrice M de u dans la base canonique de $\mathbb{K}^n$ est $\mathbb{K}$-semblable à $(C_{F_1}, C_{F_2}, \ldots, C_{F_s})$. Comme la matrice $M_{\mathbb{L}}$ de $u_{\mathbb{L}}$ dans la base canonique de $\mathbb{L}^n$ est égale à M par construction, on a que $M_{\mathbb{L}}$ est $\mathbb{L}$-semblable à $(C_{F_1}, C_{F_2}, \ldots, C_{F_s})$. Or, les F_k se divisant successivement sur $\mathbb{K}$, il en résulte que les $F_k^{\mathbb{L}}$ se divisent successivement sur $\mathbb{L}$. On a égalité par unicité dans le corollaire 5.9. □

On obtient ainsi facilement qu'être semblable ne dépend pas du choix du corps.

5.12. Corollaire. *Soit $\mathbb{L}$ une extension du corps $\mathbb{K}$ et soit A et B deux matrices de $\mathscr{M}_n(\mathbb{K})$ semblables sur $\mathbb{L}$. Alors elles sont semblables sur $\mathbb{K}$.*

Ceci généralise ce que l'on avait vu pour l'extension $\mathbb{C}$ de $\mathbb{R}$ dans la section 4.

6. Invariants de similitude et diagrammes de Young

Pour simplifier les choses, nous travaillons ici sur le corps de complexes. Nous avons vu deux invariants totaux : les diagrammes de Young associés au spectre d'une part, et les invariants de similitude d'autre part. Il est naturel de chercher à déduire les premiers des seconds et inversement.

Pour commencer, remarquons que si J_n est la matrice de Jordan (ou bloc de Jordan) de taille n, alors la matrice $J_n(\lambda) := J_n + \lambda \mathrm{I}_n$ est telle que son polynôme minimal est égal au polynôme caractéristique. La matrice $J_n(\lambda)$ est donc semblable à la matrice compagnon associée au polynôme $(X-\lambda)^n$. Ainsi, ses facteurs invariants se réduisent à $F_1 = (X-\lambda)^n$.

Afin de passer du local au global, on laisse en exercice le résultat classique suivant.

6.1. Exercice. *Soit M la matrice diagonale par blocs :*

$$M = \begin{pmatrix} A & 0 \\ 0 & A' \end{pmatrix},$$

où A et A' sont des matrices carrées. Montrer que le polynôme minimal de la matrice M est égal au PPCM des polynômes minimaux de A et A'.

On montre l'égalité par double divisibilité. Pour tout polynôme P de $\mathbb{K}[X]$, on a :

$$P(M) = \begin{pmatrix} P(A) & 0 \\ 0 & P(A') \end{pmatrix}.$$

Cela implique que si les λ_i sont des complexes distincts, alors la matrice donnée par les blocs diagonaux $\big(J_{n_i}(\lambda_i)\big)$ est telle que son polynôme minimal et son polynôme caractéristique sont égaux. Encore une fois, cette matrice est donc semblable à la matrice compagnon associée au polynôme $P := \prod_i (X-\lambda_i)^{n_i}$ et ses facteurs invariants se réduisent à $F_1 = P$.

On peut maintenant facilement voir que si la matrice A a pour spectre (λ_i) avec, associé à chaque λ_i, un diagramme de Young $\mathrm{Y}(A)_i$ possédant une suite de colonnes décroissantes $\big(d_k^*(\lambda_i)\big)$, les facteurs invariants s'obtiennent alors de la façon suivante.

Pour le premier facteur invariant, on a

$$F_1(A) = \pi_A = \prod_{\lambda_i \in \mathrm{Spec}(A)} (X-\lambda_i)^{d_1^*(\lambda_i)},$$

puisqu'il s'agit du polynôme minimal de A. Puis,

$$F_k(A) = \prod_{\lambda_i \in \mathrm{Spec}(A)} (X - \lambda_i)^{d_k^*(\lambda_i)}.$$

Nous illustrons cela sur un exemple.

6.2. Exemple. Soit

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 2 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}.$$

En développant le déterminant de $A - \lambda \mathrm{I}_5$ deux fois par rapport à la première colonne, on trouve une matrice triangulaire inférieure[6]. Il vient :

$$\chi_A(X) = -(X-2)^4(X-4) \quad \text{et} \quad \mathrm{Spec}(A) = \{2, 4\}.$$

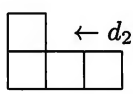
• On calcule :

$$K_1 = \mathrm{Ker}(A - 2\mathrm{I}_5) = \langle (1,0,0,0,0), (0,1,0,0,0), (0,0,0,1,0) \rangle,$$

ce qui donne $\dim K_1 = 3$. Puis,

$$K_2 = \mathrm{Ker}(A - 2\mathrm{I}_5)^2 = \langle (1,0,0,0,0), (0,1,0,0,0), (0,0,1,0,0), (0,0,0,1,0) \rangle,$$

ce qui donne $\dim K_2 = 4$.
Pas besoin d'aller plus loin car la suite (d_i) croît : $\dim K_3 = 4$.
On a la partition de 4 suivante : $d_1 = 3 \geqslant d_2 = 1$, correspondant au diagramme



En lisant le diagramme verticalement, la partition duale est $2 \geqslant 1 \geqslant 1$. Son premier élément, correspondant au nombre de cases de la première colonne, nous donne la multiplicité de la valeur propre 2 dans le polynôme minimal, à savoir 2 ici (ou encore la taille du bloc de Jordan associé à la partie nilpotente associée à la valeur propre 2).
Le deuxième nous donne la multiplicité dans un second invariant et le troisième de même (ou la taille des autres blocs de Jordan).
Nous avons donc parmi les invariants (incomplets !), $(X-2)^2$ une partie de notre polynôme minimal, et les polynômes $X-2$ et $X-2$ (ou les blocs de Jordan de taille correspondante).

[6]Noter autrement que la matrice A est triangulaire en blocs.

- Par ailleurs, on a : $K_1 = \mathrm{Ker}(A - 4\mathrm{I}_5) = \langle (0,0,0,0,1) \rangle$, d'où : $\dim K_1 = 1$ (normal pour une racine simple de χ_A !). On a une partition triviale correspondant au diagramme

Ainsi, la valeur propre 4 est de multiplicité 1 dans le polynôme minimal (ce qui n'est pas une surprise vu qu'il divise χ_A). Il faudra donc ajouter le facteur $(X-4)$ à notre polynôme minimal.

On a donc $F_1(A) = \pi_A = (X-2)^2(X-4)$ pour premier facteur invariant, les autres invariants étant $F_2(A) = X - 2$ et $F_3(A) = X - 2$.

Cela nous conduit aux formes de Jordan J_2, J_1, J_1 ajoutées à l'identité fois la valeur propre 2 et J_1 ajoutée à l'identité fois la valeur propre 4, tout cela à permutation des blocs près.

$$A \sim \begin{pmatrix} J_2 + 2\mathrm{I}_2 & & & \\ & J_1 + 2\mathrm{I}_1 & & \\ & & J_1 + 2\mathrm{I}_1 & \\ & & & J_1 + 4\mathrm{I}_1 \end{pmatrix} = \begin{pmatrix} 2 & & & & \\ & 2 & & & \\ & & 2 & 0 & \\ & & 1 & 2 & \\ & & & & 4 \end{pmatrix}.$$

Sur $\mathbb{C}$, on peut traduire facilement le corollaire 2.7.3 en terme de facteurs invariants.

6.3. Proposition. *Dans $\mathscr{M}_n(\mathbb{C})$,*

(i) *L'ensemble des matrices telles dont le facteur invariant F_1 est à racines simples est fermé.*

(ii) *L'ensemble des matrices qui n'ont qu'un facteur invariant (elles sont donc cycliques) est un ouvert.*

6.4. Question. Comment obtenir un énoncé analogue sur $\mathbb{R}$?

A. Annexe. Lemme des noyaux et décomposition de Dunford

On rappelle quelques résultats essentiels pour les théorèmes de réduction. On pourra trouver les preuves complètes dans le livre de Rémi Goblot *Algèbre linéaire* [15], chapitre 7.

Attention, pour des raisons de commodité, on désigne ici par A un endomorphisme et non pas une matrice, mais cela revient au même une fois une base fixée !

A.1. Lemme (Lemme des noyaux). *Soit A un endomorphisme d'un espace vectoriel E et $P = \prod_{i=1}^{r} P_i$ un polynôme annulateur de A dont les facteurs P_i $(1 \leqslant i \leqslant r)$ sont deux à deux premiers entre eux.*

1. *On a :*
$$E = \operatorname{Ker} P(A) = \bigoplus_i \operatorname{Ker} P_i(A).$$
2. *Pour i compris entre 1 et r, on pose $Q_i = \prod_{j \neq i} P_j$. Alors, les Q_i sont (globalement) premiers entre eux et si $\sum_i U_i Q_i = 1$ est une relation de Bézout, alors $\pi_i = U_i(A)Q_i(A)$ est le projecteur sur $\operatorname{Ker} P_i(A)$ parallèlement à $\oplus_{j \neq i} \operatorname{Ker} P_j(A)$.*
3. *En particulier, si $P = \prod_{i=1}^{r} (X - \lambda_i)^{n_i}$ est le polynôme caractéristique de A, de sorte que les λ_i soient tous distincts, alors E est la somme directe des sous-espaces caractéristiques $E^{\lambda_i} = \operatorname{Ker}(A - \lambda_i \mathrm{I}_n)^{n_i}$, qui sont stables par A.*

Esquisse. La preuve repose essentiellement sur la relation de Bézout
$$\sum_i U_i Q_i = 1,$$
que l'on applique à A pour obtenir l'identité matricielle $\sum_i U_i(A)Q_i(A) = \mathrm{Id}$. Cette relation, appliquée à un vecteur X donne la décomposition en somme. En utilisant la décomposition et l'identité $P_i(A)Q_i(A) = P(A) = 0$, on montre que la somme est directe, et le reste en découle. □

A.2. Remarque.

Les π_i sont appelés *projecteurs spectraux.* Ils forment un système de projecteurs orthogonaux, au sens où : $\sum_i \pi_i = \mathrm{Id}$ et $\pi_i \pi_j = 0$ dès que $i \neq j$.

Pour toute valeur propre λ le sous-espace E^λ est appelé *sous-espace caractéristique.* La restriction de l'endomorphisme A à E^λ est bien un endomorphisme et n'a qu'une valeur propre : λ (elle possède un polynôme annulateur ayant pour unique racine λ !). Et donc, ses propriétés sont très proches de celles d'un endomorphisme nilpotent ; deux tels endomorphismes diffèrent d'une homothétie.

Si l'on fait une analogie avec l'analyse, l'étude de restriction aux sous-espaces caractéristiques peut être vue comme une étude locale. Le lemme des noyaux peut être vu comme un passage du local au global, c'est-à-dire comme un théorème de partition de l'unité, exprimé par la formule : $\sum_i \pi_i = \mathrm{Id}$.

Si P est le polynôme caractéristique χ_A, alors en remarquant que $A\,|_{E^\lambda}$ admet pour seule valeur propre le scalaire λ et que $\chi_A = \prod_\lambda \chi_{A|_{E^\lambda}}$, on obtient $n_i = \dim\bigl(\operatorname{Ker}(A - \lambda_i \mathrm{I}_n)^{n_i}\bigr)$.

A.3. Théorème (Décomposition de Dunford). *Soit A un endomorphisme d'un espace vectoriel de dimension finie. Si son polynôme minimal π_A est scindé, alors il existe un unique couple d'endomorphismes (D, N) tel que D soit diagonalisable, N nilpotente et tel que :*

1. $A = D + N$,
2. $DN = ND$.

On a de plus :

1. *les endomorphismes D et N sont des polynômes en A ;*
2. *les endomorphismes A et D ont le même polynôme caractéristique : $\chi_D = \chi_A$.*

Esquisse. Pour l'existence, il suffit d'utiliser le lemme des noyaux et poser

$$D = \sum_{\lambda} \lambda \pi_\lambda(A), \ N = A - D,$$

où π_λ est le projecteur spectral associé à la valeur propre λ.

Alors, D est diagonalisable, D et N commutent entre eux et aussi avec A ; enfin, le polynôme caractéristique de N est X^n.

Pour l'unicité, on suppose que $D + N = D' + N'$ sont deux telles décompositions et que D et N sont les matrices définies ci-dessus. Comme D' commute avec D' et N', elle commute avec A, et donc avec D (qui est un polynôme en A). Comme D et D' commutent et qu'elles sont diagonalisables, $D - D'$ est diagonalisable (car D et D', qui commutent, le sont dans une même base). De plus, $N' - N$ est nilpotent (comme N et N' commutent, on peut appliquer le binôme de Newton), si bien que l'endomorphisme $D - D' = N' - N$, étant diagonalisable et nilpotent, est nul. □

A.4. Exemples. L'égalité

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

n'est *pas* une décomposition de Dunford, car les deux matrices du membre de droite ne commutent pas... La décomposition de Dunford est tout simplement $A = A + 0$, puisque A est diagonalisable (deux valeurs propres distinctes en dimension 2).

En revanche, l'égalité

$$B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

est bien la décomposition de Dunford de B.

A.5. Remarque. Le fait que D et N sont des polynômes en A est un élément crucial de la preuve (il permet de déduire la commutation avec D et N de la commutation avec A) et souvent utile en pratique.

De plus, si A est réelle, alors le spectre de A est invariant par l'automorphisme bar, et ceci implique que ces polynômes sont bar-invariants. Ainsi, les polynômes sont réels et donc D et N sont également réels.

A.6. Remarque. Sur $\mathbb{C}$, le polynôme minimal est toujours scindé, donc on peut toujours appliquer la décomposition de Dunford. Sur $\mathbb{R}$, on peut appliquer la décomposition de Dunford de $\mathbb{C}$, puisqu'on est plongé dans l'algèbre des matrices complexes. On a donc $A = D + N$, avec D et N complexes. Mais D et N sont réels par la remarque précédente. C'est la décomposition de Dunford dans $\mathbb{R}$. On y a gagné que l'hypothèse « polynôme minimal scindé » devient superflue, et le prix à payer est que D, qui est réelle, n'est pas diagonalisable sur $\mathbb{R}$, mais elle l'est, vue comme matrice complexe.

A.7. Remarque. Notons pour finir qu'il existe aussi une décomposition de Dunford multiplicative, plus adaptée aux groupes.

Si A est dans $\mathrm{GL}_n(\mathbb{K})$ et admet un polynôme minimal scindé, alors $A = DU$, où D et U commutent, D diagonalisable et U *unipotente*, *i.e.* $U - \mathrm{Id}$ est nilpotente. De plus, une telle décomposition est unique. La preuve est la même. Notons que, sur $\mathbb{R}$ et $\mathbb{C}$, on pourra passer de la décomposition additive à la décomposition multiplicative par l'exponentielle ; voir remarque VI-2.2.

B. Annexe. Combinatoire des partitions et diagrammes de Young

B.1. Partitions et diagrammes de Young

B.1.1. Définition. On appelle *partition* d'un entier naturel n toute suite d'entiers naturels $(\lambda_j)_{j\geqslant 1}$ qui est décroissante au sens large, nulle à partir d'un certain rang $m+1$, et dont la somme des termes vaut n : $\sum_{j=1}^{m} \lambda_j = n$. On note alors : $\lambda \vdash n$.

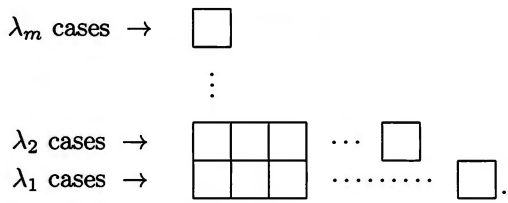
On appelle *parts* d'une partition λ les termes λ_j non nuls. Quitte à oublier ou ajouter une infinité de zéros, on peut identifier une partition à une suite finie décroissante d'entiers naturels non nuls.

À une partition λ, on associe une suite d'entiers (k_i) (croissante et qui s'essouffle) par :

$$\forall j \in \mathbb{N}^*, \quad k_j = \sum_{i=1}^{j} \lambda_i.$$

Le diagramme de Young est une visualisation d'une partition.

B.1.2. Définition (informelle). On appelle *diagramme de Young* de taille n associé à une partition λ de n est une représentation de la partition λ par n cases juxtaposées de la façon suivante :



Il y a une bijection plus ou moins claire entre partitions et diagrammes de Young. Si Y est un diagramme de Young, on notera $\lambda(Y)$ la partition associée et l'on posera, pour tout i :

$$k_i(Y) := \sum_{j=1}^{i} \lambda_j(Y).$$

Bien que l'« on voie bien » comment passer d'une partition à son diagramme de Young et inversement, étant dans une annexe, on se permet de formaliser un peu pour avoir des preuves propres.

B.1.3. Définition (formelle). On appelle *diagramme de Young « formel »* toute partie Y de $\mathbb{N}^* \times \mathbb{N}^*$ telle que pour tout $(i,j) \in Y$, on ait : $(k,\ell) \in Y$ pour tout $1 \leqslant k \leqslant i$ et tout $1 \leqslant \ell \leqslant j$. La *taille* d'un diagramme de Young est son cardinal.

Lien avec la définition précédente : on dessine une « case », c'est-à-dire un carré de côté 1, dont le coin inférieur gauche est un point de Y, vu comme partie de $\mathbb{R}^2$.

B.1.4. Lemme. *Soit n un entier naturel. Les applications suivantes sont des bijections réciproques entre partitions de n et diagrammes de Young de taille n :*
- *à une partition λ, on associe $Y_\lambda = \{(i,j) \in \mathbb{N}^* \times \mathbb{N}^* \,:\, i \leqslant \lambda_j\}$;*
- *à un diagramme de Young Y, on associe $\lambda = \lambda_Y$ où pour tout j : $\lambda_j = \max\{i \,:\, (i,j) \in Y\}$.*

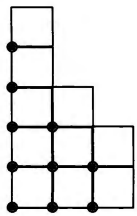


Figure B.1. *Diagramme de Young formel (points) et informel (carrés)*

B.1.5. Exemple. La partition associée au diagramme de la figure B.1 est : $\lambda = (3,3,2,1,1)$, et inversement. On comprend λ_j comme la longueur ou le nombre de cases de la ligne j du diagramme.

Démonstration. On commence par montrer que les applications sont bien définies. Soit λ une partition. Soit $(i,j) \in Y_\lambda$ et $(k,\ell) \in \mathbb{N}^2$, avec $k \leqslant i$ et $\ell \leqslant j$. Alors, puisque la suite λ est décroissante, on a : $k \leqslant i \leqslant \lambda_j \leqslant \lambda_\ell$, donc : $(k,\ell) \in Y$.

De même, soit Y un diagramme de Young et λ la suite d'entiers qui lui est associée. Soit j et ℓ deux entiers, avec $\ell \leqslant j$. On doit montrer que $\lambda_j \leqslant \lambda_\ell$. Soit $i = \lambda_j$. Alors, le point (i,j) appartient à Y, donc (i,ℓ) aussi. Par définition de λ_ℓ, on a : $i \leqslant \lambda_\ell$.

Soit λ une partition. On a : $(i,j) \in Y_\lambda$ si et seulement si $i \leqslant \lambda_j$. Pour j fixé, on peut donc écrire : $\lambda_j = \max\{m \ : \ (i,j) \in Y_\lambda\}$. Cela signifie que la partition associée à Y_λ est bien λ. Le calcul de la composée dans l'autre sens est analogue. □

B.2. Diagramme dual, partition duale

Informellement, le diagramme dual Y^* d'un diagramme de Young Y est son symétrique par rapport à la diagonale principale (figure B.2) ; la partition duale d'une partition λ s'obtient en écrivant la liste des hauteurs des colonnes du diagramme de Young de λ. Formalisons cela.

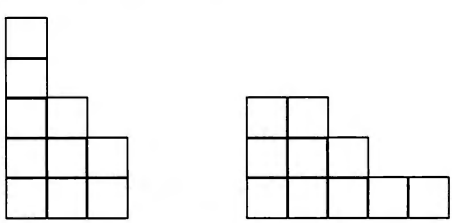


Figure B.2. *Diagrammes de Young duaux*

B.2.1. Lemme. *L'ensemble des diagrammes de Young est stable par la symétrie par rapport à la diagonale principale.*

Démonstration. Visuellement, c'est évident. Dans la version plus formelle, la symétrie est : $\tau : \mathbb{N}^2 \to \mathbb{N}^2$, $(i,j) \mapsto (j,i)$. L'image $Y^* = \tau(Y)$ d'un diagramme de Young est encore un diagramme de Young, car τ ne fait qu'échanger les conditions $k \leqslant i$ et $\ell \leqslant j$ de la définition. □

B.2.2. Lemme. *Soit $\lambda = (\lambda_j)$ une partition. Pour $i \in \mathbb{N}$, on pose :*

$$\mu_i = \#\{j \in \mathbb{N} : \ 1 \leqslant i \leqslant \lambda_j\}.$$

Alors, μ est une partition et le diagramme de Young de μ est l'image de celui de λ par la symétrie par rapport à la diagonale principale.

B.2.3. Définition. On appelle *diagramme de Young dual* d'un diagramme de Young Y et l'on note Y^* son symétrique par rapport à la diagonale principale, c'est-à-dire son image par τ.

On appelle *partition duale* d'une partition λ et l'on note λ^* la partition définie dans le lemme.

B.2.4. Exemple. Si $\lambda = (3,3,2,1,1)$, alors $\lambda^* = (5,3,2)$.

Démonstration. Il suffit de montrer que pour tout $(i,j) \in \mathbb{N}^* \times \mathbb{N}^*$, $j \leqslant \mu_i$ si et seulement si $i \leqslant \lambda_j$: en effet, cela signifie que les diagrammes de μ et λ sont symétriques par rapport à la diagonale.

Fixons i. Par définition de μ_i, il y a exactement μ_i parts λ_j de λ qui sont supérieures ou égales à i. Comme la suite λ est décroissante, ces parts sont : $\lambda_1, \ldots, \lambda_{\mu_i}$. Autrement dit : $1 \leqslant j \leqslant \mu_i$ si et seulement si $i \leqslant \lambda_j$, comme souhaité. □

B.2.5. Exercice. *Avec les notations du lemme, vérifier que μ est une partition sans passer par les diagrammes.*

B.3. Ordre de dominance

Nous allons définir de façon purement combinatoire un ordre sur les diagrammes de Young ou, ce qui revient au même, les partitions. Via la bijection entre partitions et orbites nilpotentes, cet ordre devient l'ordre de dégénérescence, ou ordre de Chevalley, sur les orbites.

B.3.1. Définition (Ordre de dominance sur les diagrammes et les partitions). Soit n un entier naturel non nul. Soit Y et Y' deux diagrammes de Young de taille n, associés aux partitions $\lambda = \lambda(Y)$ et $\lambda' = \lambda(Y')$. On dit que Y est *inférieur ou égal à Y'* et l'on note $Y \leqslant Y'$ si, pour

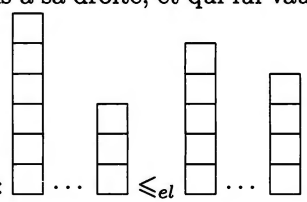
tout i, on a : $k_i(Y) \leqslant k_i(Y')$, c'est-à-dire :

$$\lambda_1 \leqslant \lambda'_1,\ \lambda_1+\lambda_2 \leqslant \lambda'_1+\lambda'_2,\ \ldots,\ \lambda_1+\cdots+\lambda_i \leqslant \lambda'_1+\cdots+\lambda'_i,\ \ldots$$

Il est immédiat de vérifier que $\leqslant$ est un ordre. On l'appelle *ordre de dominance.*

L'ordre sur l'ensemble des partitions du nombre 6 est illustré en annexe C. En général, on va donner une description plus concrète de l'ordre.

B.3.2. Définition (Relations élémentaires). Avec les mêmes notations que ci-dessus, on définit aussi la relation élémentaire $Y \leqslant_{el} Y'$ si $Y = Y'$ ou si Y' est identique à Y après qu'un bloc soit « tombé » du sommet d'une colonne sur une colonne plus à sa droite, et qui lui vaut l'appellation familière d'« ordre badaboum »[7] :



B.3.3. Proposition. *L'ordre $\leqslant$ sur les diagrammes de Young est l'ordre engendré par les relations élémentaires $\leqslant_{el}$:*

$$Y \leqslant Y' \iff \exists k \in \mathbb{N}^*,\ \exists Y_1,\ldots,Y_k,\ Y = Y_1 \leqslant_{el} Y_2 \leqslant_{el} \cdots \leqslant_{el} Y_k = Y'.$$

Démonstration. Il suffit de voir qu'entre deux diagrammes distincts Y, Y' tels que $Y \leqslant Y'$, il est possible d'insérer un diagramme Y'' distinct de Y' tel que $Y \leqslant Y'' \leqslant_{el} Y'$. À ce stade, il sera certainement utile au lecteur de suivre la preuve sur le tableau d'inégalités ci-dessous. Comme Y et Y' sont distincts, il existe r_0 minimal tel que $k_{r_0}(Y) < k_{r_0}(Y')$, donc aussi minimal tel que $\lambda_{r_0}(Y) < \lambda_{r_0}(Y')$. On note alors r l'entier maximal tel que $\lambda_r(Y') = \lambda_{r_0}(Y')$, de sorte que $\lambda_r(Y') > \lambda_{r+1}(Y')$. On a aussi par construction $k_r(Y) = \sum_{i\leqslant r}\lambda_i(Y) < \sum_{i\leqslant r}\lambda_i(Y) = k_r(Y')$. Comme $\sum_{i\leqslant n}\lambda_i(Y) = n = \sum_{i\leqslant n}\lambda_i(Y)$, il existe un $s > r$ tel que $\lambda_s(Y) > \lambda_s(Y')$, que l'on peut choisir minimal. Soit Y'' le diagramme tel que $\lambda_i(Y'') = \lambda_i(Y')$ pour $i \neq r, s$, $\lambda_r(Y'') = \lambda_r(Y') - 1$, $\lambda_s(Y'') = \lambda_s(Y') + 1$. On construit par transitivité toutes les inégalités sur les deux schémas suivants, où l'on note $\lambda_k = \lambda_k(Y)$, $\lambda'_k = \lambda_k(Y')$, $\lambda''_k = \lambda_k(Y'')$ pour tout k :

$$\begin{array}{ccccccccccccc}
\lambda'_1 & \geqslant & \cdots & \geqslant & \lambda'_{r_0-1} & \geqslant & \lambda'_{r_0} & = & \lambda'_{r_0+1} & = \cdots = & \lambda'_r & > & \lambda'_{r+1} & \geqslant & \cdots & \lambda'_{s-1} & > & \lambda'_s \\
\| & & & & \| & & \vee & & \vee & & \vee & & \vee & & & \vee & & \wedge \\
\lambda_1 & \geqslant & \cdots & \geqslant & \lambda_{r_0-1} & \geqslant & \lambda_{r_0} & \geqslant & \lambda_{r_0+1} & \geqslant \cdots \geqslant & \lambda_r & \geqslant & \lambda_{r+1} & \geqslant & \cdots & \lambda_{s-1} & \geqslant & \lambda_s.
\end{array}$$

[7] En hommage au patineur repenti Philippe Candelero.

$$\begin{array}{ccccccccccccc}
\lambda''_1 & \geqslant & \cdots & \geqslant & \lambda''_{r_0-1} & \geqslant & \lambda''_{r_0} & = & \lambda''_{r_0+1} & = \cdots \geqslant & \lambda''_r & \geqslant & \lambda''_{r+1} & \geqslant \cdots & \lambda''_{s-1} & \geqslant & \lambda''_s \\
\| & & & & \| & & \vee & & \vee & & \vee & & \vee & & \vee & & /\!\!\wedge \\
\lambda_1 & \geqslant & \cdots & \geqslant & \lambda_{r_0-1} & \geqslant & \lambda_{r_0} & \geqslant & \lambda_{r_0+1} & \geqslant \cdots \geqslant & \lambda_r & \geqslant & \lambda_{r+1} & \geqslant \cdots & \lambda_{s-1} & \geqslant & \lambda_s.
\end{array}$$

Par exemple, on a par hypothèse : $\lambda_{s-1}(Y) \geqslant \lambda_s(Y) > \lambda_s(Y') \leqslant \lambda_{s-1}(Y')$ et $\lambda_{s-1}(Y) \leqslant \lambda_{s-1}(Y')$ par minimalité de s. Donc, $\lambda_s(Y') < \lambda_{s-1}(Y')$, et l'on peut déplacer une case de la r-ième ligne de Y' en une case de la s-ième ligne de Y', et étant donné que $\lambda_r(Y') > \lambda_{r+1}(Y)$ et $\lambda_s(Y') > \lambda_{s-1}(Y')$, on obtient que Y'' est encore un diagramme de Young, et qu'il vérifie $Y'' \leqslant_{el} Y'$ par construction. De plus, les inégalités ci-dessus indiquent que $Y \leqslant Y''$: pour le voir, il suffit de se convaincre que $k_i(Y) \leqslant k_i(Y'')$ pour $i = r, s$ (les autres inégalités découlent de la construction). D'une part, on a clairement en additionnant les inégalités membre à membre que $k_r(Y) \leqslant k_r(Y'')$. D'autre part, $k_s(Y) \leqslant k_s(Y') = k_s(Y'')$, la dernière inégalité provenant de l'élimination $-1 + 1 = 0$ dans la somme exprimant l'entier $k_s(Y'')$. □

B.3.4. Exercice. *Soit Y et Z deux diagrammes de Young. Montrer l'équivalence : $Y \leqslant Z$ si et seulement si $Z^* \leqslant Y^*$.*

C. Développement. Diagrammes de Young et invariants de similitude des matrices nilpotentes 6×6

Ici, on prend $n = 6$.

Nous donnons le *diagramme de Hasse*[8] de l'ordre des diagrammes de Young de taille 6 : deux diagrammes sont reliés par une arête si l'un (le plus haut sur le graphe) est élémentairement plus petit que l'autre. On en déduit donc qu'un diagramme est plus grand que l'autre si l'on peut joindre le premier par une suite d'arêtes en allant de bas en haut sur le graphe. Par exemple, on voit l'orbite ouverte tout en haut du graphe et l'orbite nulle tout en bas.

Il est intéressant de voir que la dualité des partitions se réalise en une symétrie du graphe : voir l'exercice B.3.4.

Le polynôme caractéristique de A est le même pour tous : $\chi_A(X) = X^6$.

On note IS l'ensemble des invariants de similitude de la matrice A, lesquels ont été définis au §5. Notons juste qu'il s'agit de polynômes unitaires $F_1, \ldots, F_s$ qui ne dépendent que de la classe de conjugaison de la matrice concernée, et même qui la caractérisent.

[8] Collision malheureuse du vocabulaire.

C.1. Remarque. Les exposants de X dans les polynômes invariants sont aussi les tailles des blocs de Jordan correspondants.

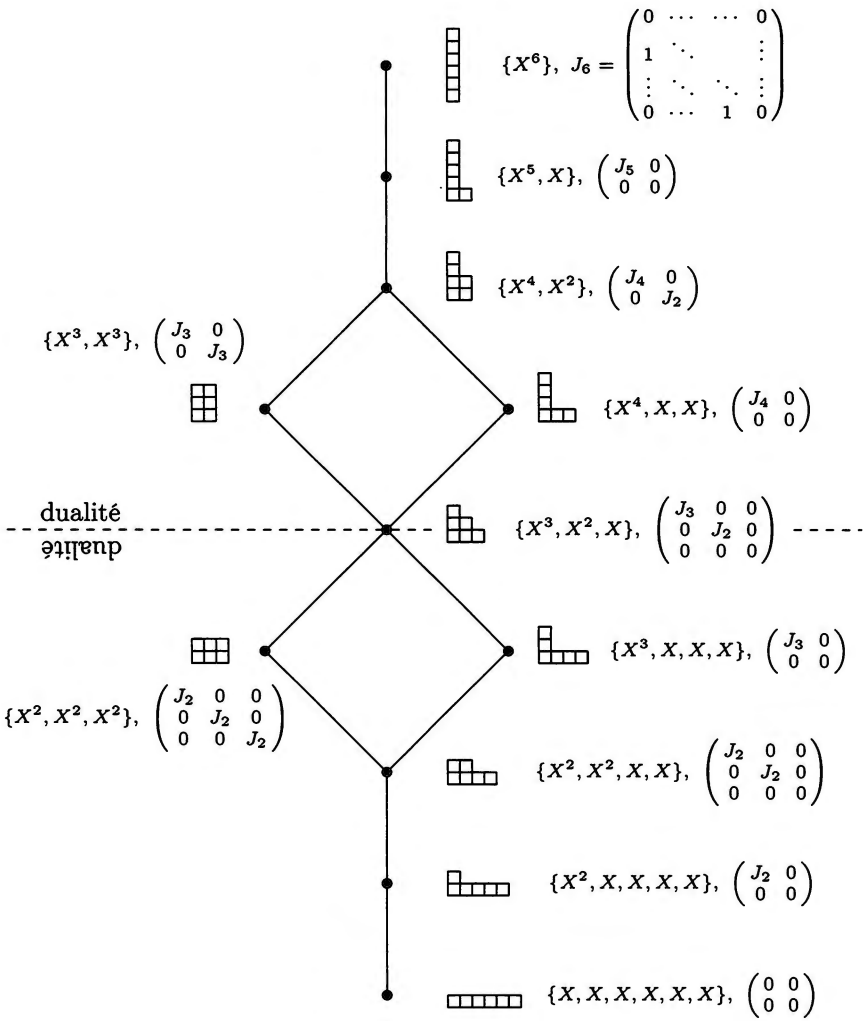


Figure C.1. *Orbites nilpotentes dans $\mathcal{M}_6(\mathbb{C})$: diagramme de Young, invariants de similitudes et formes normales*

D. Exercices

Diagonalisation, trigonalisation

D.1. Exercice (Diagonalisabilité et passage à une racine k-ème)

Soit A une matrice inversible complexe et k un entier tels que A^k est diagonalisable.

1. *Montrer que A est diagonalisable.*

   On pourra par exemple décomposer l'espace en somme de sous-espaces propres pour A^k puis appliquer le lemme des noyaux à $\mathrm{Ker}(A^k - \lambda\,\mathrm{Id})$, sachant que si λ est non nul, le polynôme $X^k - \lambda$ n'a que des racines simples.

2. *Donner un contre-exemple si l'on ne suppose pas A inversible.*

   C'est le moment de penser aux nilpotentes.

D.2. Exercice (Polynôme minimal d'une matrice diagonale par blocs).

Soit M la matrice diagonale par blocs :

$$M = \begin{pmatrix} A & 0 \\ 0 & A' \end{pmatrix},$$

où A et A' sont des matrices carrées. Montrer que le polynôme minimal de M est le PPCM des polynômes minimaux de A et A'. En déduire que M est diagonalisable si et seulement si A et A' le sont.

D.3. Exercice (Similitude et équivalence pour les projections)

Montrer que deux matrices de projections sont semblables si et seulement si elles sont équivalentes.

D.4. Exercice (Diagonalisabilité et semi-simplicité)

Sur $\mathbb{C}$, montrer que u est diagonalisable si et seulement si tout sous-espace u-stable possède un supplémentaire u-stable. Montrer que ce résultat est faux sur $\mathbb{R}$. Comment l'adapter sur $\mathbb{R}$?

Pour le cas réel, on pourra considérer des rotations du plan, qui sont d'ailleurs diagonalisables sur $\mathbb{C}$...

D.5. Exercice (Diagonalisabilité sur un corps fini)

Sur le corps $\mathbb{F}_q$, montrer qu'une matrice carrée A est diagonalisable si et seulement si $A^q - A = 0$.

Pour le « seulement si », on diagonalise et l'on utilise le fait que tout élément de $\mathbb{F}_q$ est racine de $X^q - X = 0$. Pour le « si », on montre que le polynôme $X^q - X$ est scindé à racines simples sur $\mathbb{F}_q$.

D.6. Exercice (Normalisateur des matrices diagonales inversibles)
Le sous-groupe D du groupe des matrices inversibles $n\times n$ formé des matrices diagonales est-il distingué ? Soit N son normalisateur, montrer que $N/D \simeq \mathfrak{S}_n$, puis que $N \simeq D \rtimes \mathfrak{S}_n$.
On pourra dans un premier temps considérer l'action par conjugaison du groupe des matrices de permutation sur D.

D.7. Exercice (Adhérence de l'ensemble des matrices réelles diagonalisables)

1. *Trouver une matrice de $\mathscr{M}_2(\mathbb{R})$ qui n'est pas dans l'adhérence de l'ensemble des matrices diagonalisables.*
   On pourra utiliser l'application qui à une matrice A associe le discriminant de son polynôme caractéristique.
2. *Soit (D_k) une suite de matrices diagonalisables qui converge vers A : montrer que le polynôme caractéristique de A est scindé.*
   L'exercice II-F.33 peut être utile.
3. *Montrer que toute matrice trigonalisable est la limite d'une suite de matrices diagonalisables.*
   Pour une matrice triangulaire, la matrice diagonale dont les coefficients non nuls sont $(1, t, t^2, \dots)$ est ton amie !
4. *Conclure : l'adhérence de l'ensemble des matrices réelles diagonalisables est l'ensemble des matrices trigonalisables.*

D.8. Exercice. *Montrer que si un ouvert de $\mathscr{M}_n(\mathbb{C})$ contient les matrices diagonales et est stable par conjugaison, alors il est égal à $\mathscr{M}_n(\mathbb{C})$ tout entier.*
On pourra montrer, *a contrario*, que l'adhérence de la classe de similitude de toute matrice contient une matrice diagonale. La matrice diagonale dont les coefficients non nuls sont $(1, t, t^2, \dots)$ est ton amie.

D.9. Exercice (Classes de conjugaison bornées)
Montrer que les seules classes de conjugaison bornées dans $\mathscr{M}_n(\mathbb{C})$ (pour sa structure d'espace vectoriel normé) sont les classes des homothéties.
Regarder d'abord le cas de la dimension 2.

D.10. Exercice (Intérieur de l'ensemble des matrices diagonalisables)

1. *Montrer que l'ensemble des matrices diagonalisables complexes à valeurs propres distinctes est un ouvert.*
   On pourra utiliser l'application qui à A associe le discriminant du polynôme caractéristique de A.
2. *Montrer qu'il s'agit de l'intérieur de l'ensemble des matrices diagonalisables.*

Lemme des noyaux, Dunford et réduite de Jordan

D.11. Exercice (Variations sur le lemme des noyaux)
Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel (supposé, par prudence, de dimension finie) et soit P et Q deux polynômes de $\mathbb{K}[X]$.

1. *Montrer les deux égalités qui suivent :*

$$\operatorname{Ker}\operatorname{pgcd}(P,Q)(u) = \operatorname{Ker}P(u) \cap \operatorname{Ker}Q(u)$$
$$\operatorname{Im}\operatorname{pgcd}(P,Q)(u) = \operatorname{Im}P(u) + \operatorname{Im}Q(u).$$

2. *Montrer également que*

$$\operatorname{Ker}\operatorname{ppcm}(P,Q)(u) = \operatorname{Ker}P(u) + \operatorname{Ker}Q(u)$$
$$\operatorname{Im}\operatorname{ppcm}(P,Q)(u) = \operatorname{Im}P(u) \cap \operatorname{Im}Q(u).$$

On pourra penser à l'identité de Bézout, au lemme des noyaux et, pour le passage du noyau à l'image, à la dualité.

D.12. Exercice (Décomposition de Dunford d'un crochet).
Soit A une matrice complexe carrée. On note

$$\varphi_A : \mathscr{M}_n(\mathbb{C}) \longrightarrow \mathscr{M}_n(\mathbb{C}), \quad M \longmapsto AM - MA.$$

1. *Montrer que si A est diagonalisable, l'endomorphisme φ_A l'est aussi.*
   Commencer par le cas d'une matrice diagonale et exhiber une base propre. Ramener le cas général à ce cas.
2. *Montrer que si A est nilpotente, φ_A l'est aussi.*
   On calcule φ_A^k en remarquant que φ_A est la somme de deux endomorphismes nilpotents qui commutent : la multiplication à gauche par A et la multiplication à droite par $-A$.
3. *En déduire la décomposition de Dunford de φ_A dans* $\operatorname{End}\big(\mathscr{M}_n(\mathbb{C})\big)$.

D.13. Exercice (Cas particulier du théorème de Jordan-Kronecker)
On fait agir le groupe $G = \mathrm{GL}_n(\mathbb{C}) \times \mathrm{GL}_n(\mathbb{C})$ sur lui-même par

$$(P,Q)\cdot(A,B) = (PAQ^{-1}, PBQ^{-1}).$$

(Attention, il ne s'agit pas de l'action de G sur lui-même par conjugaison.)

1. *Montrer que chaque orbite contient un couple de matrices de la forme (I_n, J), où J est une réduite de Jordan.*
   On pourra dans l'orbite de (A,B) considérer la réduite de Jordan associée à $A^{-1}B$.
2. *Montrer inversement que si (I_n, J) et (I_n, J') sont dans la même orbite, où J et J' sont des réduites de Jordan, alors J et J' sont égales à permutation des blocs près.*
3. *En déduire un invariant total pour cette action.*

D.14. Exercice (Non-continuité de la décomposition de Dunford)
Montrer que dans la décomposition de Dunford d'une matrice complexe, disons $M = D + N$ avec D est diagonalisable et N nilpotente, l'application $M \mapsto N$ n'est pas continue. Est-ce que l'application $M \mapsto D$ l'est ?
Elle est constante sur les matrices diagonalisables, et les matrices diagonalisables sont denses.

Nilpotence

D.15. Exercice (Commutant d'une matrice de Jordan)
On rappelle que J_n est la matrice de Jordan indécomposable $n \times n$. Soit A une matrice qui commute à J_n. Soit (e_i) la base canonique de $\mathbb{K}^n$.

1. *Remarquer que $J_n^k e_1 = e_{1+k}$ pour k de 0 à $n-1$. En déduire que Ae_1 peut s'écrire $P(J_n)e_1$, où P est un polynôme de degré inférieur ou égal à $n-1$.*
2. *Montrer que $Ae_i = P(J_n)e_i$ pour tout i.*
3. *Montrer que le commutant de J_n est égal au sous-espace des polynômes en cette matrice.*
4. *Quel est le nombre de matrices nilpotentes de taille $n \times n$ à coefficients dans $\mathbb{F}_q$ et de rang $n-1$?*
   Elles sont toutes $\mathrm{GL}_n(\mathbb{F}_q)$-semblables à J_n.

D.16. Exercice (Familles de matrices nilpotentes qui commutent)
Soit $(f_1, \ldots, f_n)$ une famille d'endomorphismes nilpotents de $\mathcal{M}_n(\mathbb{K})$ qui commutent deux à deux. Montrer que $f_1 \circ f_2 \circ \cdots \circ f_n = 0$.
On pourra utiliser la trigonalisation simultanée.

D.17. Exercice (Un critère de nilpotence)
Montrer que A est nilpotente si et seulement si A est semblable à $2A$.
Pour l'implication directe, on pourra utiliser des réduites de Jordan. Pour la réciproque, on pourra remarquer que si la matrice nulle est dans l'adhérence de la classe de conjugaison de A, alors A est nilpotente.

D.18. Exercice. *Montrer que si A commute avec $AB-BA$, alors $AB-BA$ est nilpotente.*
On montre (sans user d'une récurrence) que la trace de $(AB-BA)^m$ est nulle pour tout $m > 0$. Pour cela, on constate que l'on a :

$$\begin{aligned}\mathrm{tr}(AB-BA)^m &= \mathrm{tr}(AB(AB-BA)^{m-1} - BA(AB-BA)^{m-1}) \\ &= \mathrm{tr}\left(AB(AB-BA)^{m-1} - B(AB-BA)^{m-1}A\right) \\ &= \mathrm{tr}\left(AB(AB-BA)^{m-1} - (AB-BA)^{m-1}AB\right) = 0.\end{aligned}$$

D.19. Exercice (Radical nilpotent de l'annulateur)

1. *Soit A une matrice carrée à coefficients complexes. Montrer que l'ensemble des polynômes P tels que $P(A)$ est nilpotente est un idéal de l'anneau principal $\mathbb{C}[X]$ et en donner un générateur.*

   L'idéal est engendré par $\prod_{\lambda \in \mathrm{Spec}(A)}(X-\lambda)$. C'est ce que l'on appelle le radical (nilpotent) de l'idéal annulateur de A.

2. *Quels sont les éléments nilpotents de $\mathbb{Z}/48\mathbb{Z}$? de $\mathbb{Z}/n\mathbb{Z}$, pour n entier quelconque ? Quels sont les entiers a qui définissent un endomorphisme nilpotent du $\mathbb{Z}$-module $\mathbb{Z}/n\mathbb{Z}$? Quel rapport avec la question précédente ?*

Instructif

D.20. Exercice (Un isomorphisme de Harish-Chandra)

On considère la $\mathbb{C}$-algèbre $\mathbb{C}[\mathscr{M}_n(\mathbb{C})]$ des fonctions polynomiales sur $\mathscr{M}_n(\mathbb{C})$, identifié à l'algèbre de polynômes sur $\mathbb{C}$ à n^2 indéterminées. La $\mathbb{C}$-algèbre

$$\left\{f \in \mathbb{C}[\mathscr{M}_n(\mathbb{C})],\ \forall A \in \mathscr{M}_n(\mathbb{C}), \forall P \in \mathrm{GL}_n(\mathbb{C}),\ f(PAP^{-1}) = f(A)\right\}$$

sera notée $\mathbb{C}[\mathscr{M}_n(\mathbb{C})]^{\mathrm{GL}_n(\mathbb{C})}$.

1. *Vérifier qu'il s'agit bien d'une $\mathbb{C}$-algèbre.*

2. *Montrer que la restriction au sous-espace $\mathscr{D}_n$ des matrices diagonales définit un morphisme*

$$\rho : \mathbb{C}[\mathscr{M}_n(\mathbb{C})]^{\mathrm{GL}_n(\mathbb{C})} \longrightarrow \mathbb{C}[\mathscr{D}_n]^{\mathfrak{S}_n},$$

   où $\mathbb{C}[\mathscr{D}_n]^{\mathfrak{S}_n}$ est la $\mathbb{C}$-algèbre des fonctions polynomiales symétriques sur $\mathscr{D}_n$ identifiée à l'algèbre des polynômes symétriques à n indéterminées.

3. *En utilisant le fait que les matrices diagonalisables sont denses dans $\mathscr{M}_n(\mathbb{C})$, montrer que ρ est injective.*

4. *En utilisant le fait que les polynômes symétriques élémentaires engendrent l'algèbre $\mathbb{C}[\mathscr{D}_n]^{\mathfrak{S}_n}$, montrer que ρ est un isomorphisme.*

D.21. Exercice. *Montrer (sans utiliser les invariants de similitude) que la dimension du commutant d'un endomorphisme ne varie pas par extension du corps.*

On rappelle que le rang d'une matrice ne varie pas par extension de corps...

D.22. Exercice (« Désingularisation » de Springer)

Soit $E = \mathbb{C}^n$. On considère dans $\mathrm{End}(E)$ le cône $\mathscr{N}$ des éléments nilpotents de $\mathrm{End}(E)$, et la variété de drapeaux $\mathscr{F}$ de E.

$$\mathscr{N} := \{x \in \mathrm{End}(E),\ x^n = 0\},\ \mathscr{F} := \{F_0 \subset F_1 \subset \cdots \subset F_n,\ \dim F_i = i\}.$$

1. *Montrer que le groupe* $G = \mathrm{GL}_n(\mathbb{C})$ *agit transitivement sur* $\mathscr{F}$ *par* $g\cdot(F_i) = (gF_i)$. *Montrer que le stabilisateur d'un élément est isomorphe au sous-groupe constitué des matrices* $n\times n$ *inversibles triangulaires supérieures.*
2. *En déduire que l'on peut munir* $\mathscr{F}$ *d'une topologie telle que l'action de* G *sur* $\mathscr{F}$ *est une action continue. En étudiant la restriction de cette action au sous-groupe* SU_n, *montrer que* $\mathscr{F}$ *est compact.*
3. *On fixe un drapeau* $F_* = (F_i)$ *de* $\mathscr{F}$. *Soit* $\mathcal{N}^{F_*}$ *l'ensemble des éléments* x *de* $\mathcal{N}$ *tels que* x *stabilise* F_*, i.e. $x(F_i) \subset F_i$ *pour tout* i.
   (a) *Montrer que pour tout* x *de* $\mathcal{N}^{F_*}$ *et pour tout* i, *on a :* $x(F_i) \subset F_{i-1}$.
   (b) *En déduire que* $\mathcal{N}^{F_*}$ *est un sous-espace vectoriel de* $\mathrm{End}(E)$.
   (c) *Soit* $g \in G$. *Montrer que si* x *est dans* $\mathcal{N}^{F_*}$, *alors la matrice* gxg^{-1} *est dans* $\mathcal{N}^{gF_*}$.
   (d) *En déduire que la dimension de l'espace vectoriel* $\mathcal{N}^{F_*}$ *ne dépend pas du drapeau* F_* *choisi.*
4. *Soit* x *dans* $\mathcal{N}$. *On considère l'ensemble* $\mathscr{F}^x$ *des drapeaux que* x *stabilise. Montrer que* $\mathscr{F}^x$ *est compact.*
5. *Décrire* $\mathscr{F}^x$:
   (a) *lorsque* x *est nul ;*
   (b) *lorsque* x *admet pour matrice un bloc de Jordan indécomposable.*

D.23. Exercice (Deux endomorphismes qui commutent)
Soit V *un endomorphisme de dimension finie* n.

1. *Soit* v *un endomorphisme de* V. *Pour tout vecteur* x *de* V, *on note* I_x *l'ensemble des polynômes* P *de* $\mathbb{C}[X]$ *tels que* $P(v)(x) = 0$. *Vérifier que* I_x *est un idéal de* $\mathbb{C}[X]$ *non réduit à* $\{0\}$, *et que son unique générateur unitaire* μ_x *divise le polynôme minimal de* v. *Pour un vecteur* x *de* V, *prouver l'équivalence des conditions suivantes :*
   (i) *le polynôme* μ_x *est de degré* n ;
   (ii) *les vecteurs* $x, v(x), \ldots, v^{n-1}(x)$ *sont linéairement indépendants.*
   On dit que v *est* régulier *s'il existe un vecteur* x *de* V *vérifiant ces conditions.*
2. *Soit* v *un endomorphisme régulier de* V. *Prouver que les endomorphismes qui commutent à* v *sont les polynômes en* v.
3. *Soit* v *un endomorphisme de* V *et soit* w *un endomorphisme* régulier *de* V. *On fixe un vecteur* x *de* V *tel que* $\big(x, w(x), \ldots, w^{n-1}(x)\big)$ *soit une base de* V, *que l'on notera* $\mathscr{B}$. *Prouver que* $v + \epsilon w$ *est régulier pour tous les nombres complexes* ϵ *sauf peut-être un nombre fini d'entre eux.*

4. *Soit v un endomorphisme de V. On suppose que dans une base $\mathscr{B}$ de V, l'endomorphisme v admet pour matrice la réduite de Jordan J_λ pour une partition $\lambda = (\lambda_1, \dots, \lambda_\alpha)$ de n avec α parts non nulles. Soit $c_1, \dots, c_\alpha$ des nombres complexes* distincts *deux à deux, et soit w l'endomorphisme de V ayant pour matrice dans la base $\mathscr{B}$ la matrice diagonale par blocs dont les blocs successifs sont $c_1 \mathrm{I}_{\lambda_1}, \dots, c_\alpha \mathrm{I}_{\lambda_\alpha}$. Prouver que $v + w$ est régulier.*

5. *Soit u un endomorphisme* nilpotent *de V. Prouver qu'il existe un endomorphisme régulier w qui commute à u.*

6. *Soit u un endomorphisme nilpotent de V et soit v un endomorphisme de V qui commute avec u. On note A le sous-espace vectoriel de* $\mathrm{End}(V)$ *engendré par les endomorphismes $u^i v^j$ lorsque i et j parcourent les entiers naturels. Prouver que A est de dimension au plus n.*

   On pourra traiter d'abord le cas où v est régulier et utiliser le fait que le cône nilpotent est fermé pour le cas général.

7. *Soit u et v deux endomorphismes nilpotents de V qui commutent entre eux. On note B le sous-espace vectoriel de* $\mathrm{End}(V)$ *engendré par les endomorphismes $u^i v^j$ quand (i, j) parcourt les couples d'entiers naturels non tous deux nuls. Prouver que B est de dimension au plus $n - 1$.*

8. *Pour $n = 4$, donner un exemple d'endomorphismes u et v comme dans la question 7, tels que B soit de dimension 3 mais ne contienne aucun endomorphisme régulier.*

   On pourra chercher des endomorphismes dont la matrice dans une base donnée de V est triangulaire supérieure, et à coefficients 0 ou 1.

Cet exercice est une partie de l'épreuve d'entrée à l'École normale supérieure de 2011.

Chapitre IV

Matrices échelonnées et grassmanniennes

Wong Fei-hung : [Drinking some very strong alcohol in the middle of a fight] What the hell is that ?
Mrs. Wong : What does it mean when there's a picture of a skull ?
Wong Fei-hung : Good stuff!!!
Legend of the Drunken Master, Chia-Liang Liu, 1994.

On a déjà vu au chapitre I comment démontrer le théorème du rang à l'aide d'opérations élémentaires sur les rangées (proposition I-B.1). Plus précisément, en multipliant à gauche et à droite une matrice par des matrices élémentaires, on a ramené toute matrice à sa « forme normale », c'est-à-dire à une matrice $\mathrm{I}_{m,n,r}$. On a interprété cette procédure comme la recherche d'un invariant (le rang) pour l'action du produit de groupes linéaires $\mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$ sur l'espace des matrices $\mathscr{M}_{m,n}(\mathbb{K})$.

Or, dans certaines situations, la plus typique étant la résolution de systèmes linéaires, il est très raisonnable de faire des opérations sur les lignes et un peu maladroit de manipuler les colonnes (cela se traduirait par des changements de variables portant sur les inconnues). Se pose donc naturellement la question de trouver une forme normale des matrices rectangulaires de $\mathscr{M}_{m,n}(\mathbb{K})$ sous l'action d'un groupe linéaire $\mathrm{GL}_m(\mathbb{K})$ par multiplication à gauche. C'est ainsi qu'apparaissent, très naturellement, les matrices échelonnées.

En transposant la situation, c'est-à-dire en manipulant les colonnes des matrices, on peut l'interpréter de façon géométrique comme une sorte de classification des sous-espaces d'un espace vectoriel. Il s'agit d'étudier la

position du sous-espace par rapport à un drapeau de référence et, encore une fois, on reformulera le problème comme l'étude des orbites d'un sous-groupe du groupe linéaire dans un quotient (« les B-orbites dans G/P », notations à venir). Cette décomposition sera exploitée au tome suivant, § IV-1 dans le cas d'un corps fini pour donner de très jolies formules combinatoires (q-binôme de Newton, triple produit de Jacobi et applications).

Il sera alors naturel d'étendre ces considérations des sous-espaces aux drapeaux (des suites de sous-espaces emboîtés), ce qui donne lieu à la *décomposition de Bruhat*[1]. Cela sera fait dans le second volume.

Il est remarquable que toutes ces considérations soient essentiellement des reformulations, en termes géométriques ou de théorie des groupes, de variantes de l'algorithme du pivot de Gauss.

Dans tout le chapitre, $\mathbb{K}$ désigne un corps, qui peut être quelconque tant qu'il n'est pas question de topologie, le corps des réels ou des complexes lorsque l'on parle d'adhérence et de choses afférentes.

1. Motivation

Systèmes linéaires

Lorsque l'on résout un système linéaire de la forme $AX = Y$, où A est une matrice rectangulaire à m lignes et n colonnes et Y un vecteur-colonne à m coefficients sont donnés et où X est un vecteur-colonne inconnu à n lignes, l'algorithme du pivot de Gauss donne une méthode pour simplifier le système en un système qui lui est équivalent par des opérations élémentaires sur les équations. Matriciellement, cela revient à multiplier A et Y par une même matrice $m\times m$ *inversible* P et à utiliser l'équivalence :

$$AX = Y \iff PAX = PY.$$

Rappelons que les opérations élémentaires s'interprètent de la façon suivante (on trouvera plus de détails en annexe A) :
- multiplier une équation par une constante revient à multiplier la matrice à gauche par une matrice de dilatation ;
- faire une combinaison linéaire de lignes d'une matrice revient à la multiplier à gauche par une matrice de transvection ;
- permuter deux lignes revient à multiplier à gauche par une matrice de permutation ;
- enfin, le groupe linéaire est engendré par les dilatations et les transvections.

Ainsi, on cherche P de sorte que PA soit aussi simple que possible.

[1] François Bruhat (1929-2007), mathématicien français spécialiste des groupes algébriques et de la théorie géométrique des groupes.

1.1. Problème. Existe-t-il une forme normale qui caractérise les orbites de l'action de $\mathrm{GL}_m(\mathbb{K})$ sur $\mathscr{M}_{m,n}(\mathbb{K})$ par multiplication à gauche,

$$(P, A) \mapsto P \cdot A = PA\,?$$

Les matrices échelonnées réduites répondent positivement au problème.

Grassmanniennes

Soit n un entier naturel et m un entier compris entre 1 et n. On cherche un paramétrage de la grassmannienne $\mathrm{Gr}_{m,n}(\mathbb{K})$ des sous-espaces de dimension m de $\mathbb{K}^n$. Un tel sous-espace F possède une base $\mathbf{f} = (f_1, \ldots, f_m)$: si l'on écrit la matrice dont les colonnes sont les coordonnées des vecteurs de $\mathbf{f}$ dans la base standard de $\mathbb{K}^n$, on décrit F par une matrice $n \times m$ de rang m. Mais cette description n'est pas canonique puisqu'il faut choisir une base de F.

On obtient en tout état de cause une surjection $\varphi : \mathscr{M}^0_{n,m}(\mathbb{K}) \to \mathrm{Gr}_{m,n}(\mathbb{K})$ définie sur l'ouvert $\mathscr{M}^0_{n,m}$ des matrices $n \times m$ de rang m : à une matrice A, on associe le sous-espace de $\mathbb{K}^n$ engendré par les colonnes de A.

1.2. Lemme. *Soit n et m deux entiers, avec $m \leqslant n$. Soit A et A' deux matrices $n \times m$ de rang m. Il est équivalent de dire :*

(i) *les colonnes de A et A' engendrent le même sous-espace ;*

(ii) *il existe une matrice $m \times m$ inversible Q telle que $A' = AQ$.*

Démonstration. Si les colonnes de A et A' engendrent le même sous-espace F de $\mathbb{K}^n$, l'hypothèse sur le rang implique que les colonnes de ces deux matrices forment des bases $\mathbf{f} = (f_1, \ldots, f_m)$ et $\mathbf{f}' = (f'_1, \ldots, f'_m)$ de F. Mais, si $P = (p_{ij})_{1 \leqslant i,j \leqslant m}$ désigne la matrice de passage de $\mathbf{f}$ à $\mathbf{f}'$, c'est-à-dire si

$$\forall j \in \{1, \ldots, m\}, \quad f'_j = \sum_{i=1}^{m} p_{ij} f_i,$$

on a alors : $A' = AQ$, où $Q = {}^tP$. (C'est la formule duale, moins usitée, du classique $X = PX'$.)

Inversement, si l'on a $A' = AQ$ pour une matrice Q inversible, alors les colonnes de A' sont des combinaisons linéaires des colonnes de A, ce qui donne l'inclusion d'un sous-espace dans l'autre ; l'inclusion inverse vient de même de l'égalité $A = A'Q^{-1}$. Cela prouve le lemme. □

Ainsi, par passage au quotient, φ induit une bijection :

$$\overline{\varphi} : \mathscr{M}^0_{n,m} / \mathrm{GL}_m(\mathbb{K}) \longrightarrow \mathrm{Gr}_{m,n}(\mathbb{K}),$$

où l'ensemble de départ de $\overline{\varphi}$ est l'ensemble des orbites pour l'action (à droite) de $\mathrm{GL}_m(\mathbb{K})$ par multiplication à droite sur les matrices $n \times m$.

L'étude des grassmanniennes[2] serait considérablement simplifiée si l'on disposait d'une « forme normale » pour les (matrices des) bases des sous-espaces de dimension m de $\mathbb{K}^n$.

1.3. Problème (dual du précédent). Peut-on trouver une forme simple qui caractérise chaque orbite de l'action de $\mathrm{GL}_m(\mathbb{K})$ par multiplication à droite sur les matrices $n \times m$ (de rang maximal) ?

Les deux problèmes ci-dessus sont duaux l'un de l'autre ou, plus précisément, reliés l'un à l'autre par la transposition.

2. Matrices échelonnées en lignes

Montrons que les matrices échelonnées en lignes donnent une forme normale pour les matrices rectangulaires à multiplication à gauche par une matrice inversible près.

2.1. Définition. On appelle *pivot* d'une ligne non nulle le coefficient non nul situé dans la colonne le plus à gauche. Une matrice est dite *échelonnée en lignes* lorsqu'elle vérifie les conditions suivantes :
- si une ligne est nulle, toutes les lignes suivantes sont nulles ;
- le pivot d'une ligne est strictement plus à droite que les pivots des lignes précédentes.

Une matrice échelonnée est dite *réduite* si, de plus,
- tous les pivots sont égaux à 1 et les pivots sont les seuls coefficients non nuls de leur colonne.

On appelle *type* d'une matrice échelonnée la liste strictement croissante $\mathbf{j} = (j_1, j_2, \ldots, j_r)$ des indices des colonnes des pivots, c'est-à-dire la partie $\{j_1, j_2, \ldots, j_r\}$.

Une matrice est dite *échelonnée en colonnes (réduite)* si sa transposée est échelonnée en lignes (réduite).

2.2. Exemple. Voici la forme d'une matrice échelonnée en lignes réduite de type $(2,3,5)$:

$$\begin{pmatrix} 0 & 1 & 0 & * & 0 & * & * \\ 0 & 0 & 1 & * & 0 & * & * \\ 0 & 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

où les coefficients « $*$ » sont arbitraires.

[2] Il n'est pas nécessairement opportun de parler de classification car, à multiplication à gauche par $\mathrm{GL}_n(\mathbb{K})$ près, il n'y a qu'une seule classe... En effet, la grassmannienne est un espace homogène du groupe $\mathrm{GL}_n(\mathbb{K})$.

2.3. Théorème. *Soit m et n deux entiers. On considère l'action de $\mathrm{GL}_m(\mathbb{K})$ par multiplication à gauche sur l'espace $\mathscr{M}_{m,n}(\mathbb{K})$ des matrices $m \times n$. Alors :*

(i) *deux matrices A et A' de $\mathscr{M}_{m,n}(\mathbb{K})$ sont dans la même orbite sous $\mathrm{GL}_m(\mathbb{K})$ si et seulement si elles ont le même noyau :*

$$\exists\, P \in \mathrm{GL}_m(\mathbb{K}),\ A' = PA \iff \mathrm{Ker}\, A = \mathrm{Ker}\, A' ;$$

(ii) *toute matrice est dans l'orbite d'une unique matrice échelonnée en lignes réduite :*

$$\mathscr{M}_{m,n}(\mathbb{K}) = \bigcup_{E \in \mathscr{E}_{m,n}} \mathrm{GL}_m(\mathbb{K}) \cdot E,$$

où la réunion est disjointe et où $\mathscr{E}_{m,n}$ désigne l'ensemble des matrices échelonnées en lignes réduites de taille $m \times n$.

Démonstration. (i) Si $A' = PA$ pour P inversible, le noyau de A' contient celui de A ; l'inclusion inverse provient de l'égalité $A = P^{-1}A'$.

La réciproque a été implicitement vue dans la preuve du théorème du rang ; mais on va l'expliciter. Soit φ_A et $\varphi_{A'}$ les morphismes de $\mathbb{K}^n$ vers $\mathbb{K}^m$ définis par la multiplication par A et A', respectivement. Par hypothèse, ces morphismes ont le même noyau K ; ils ont donc le même rang. Fixons une base $\mathbf{v} = (v_1, \ldots, v_r)$ d'un supplémentaire de K. Par la version abstraite du théorème du rang, $\varphi_A(\mathbf{v})$ et $\varphi_{A'}(\mathbf{v})$ sont deux familles libres de même cardinal. On les complète en deux bases $\mathbf{w} = \big(\varphi_A(v_1), \ldots, \varphi_A(v_r), v_{r+1}, \ldots, v_n\big)$ et $\mathbf{w}' = \big(\varphi_{A'}(v_1), \ldots, \varphi_{A'}(v_r), v'_{r+1}, \ldots, v'_n\big)$. L'endomorphisme ψ qui envoie $\mathbf{w}$ sur $\mathbf{w}'$ est inversible et, par construction, $\psi\varphi_A$ et $\varphi_{A'}$ coïncident sur $\mathbf{v}$. Comme $\psi\varphi_A$ et $\varphi_{A'}$ coïncident également sur K, ces applications linéaires sont égales. La matrice P de ψ dans la base canonique satisfait à l'égalité $PA = A'$.

(ii) Montrons que toute matrice est dans l'orbite d'une matrice échelonnée. C'est exactement la méthode du pivot de Gauss ! Avec des opérations élémentaires sur les lignes de la matrice donnée A, que l'on peut voir comme autant de multiplications à gauche par des matrices inversibles (voir l'annexe A), on transforme d'abord la matrice en une matrice échelonnée en lignes, puis en une matrice échelonnée en lignes réduite en annulant les coefficients éventuels au-dessus des pivots. On trouve ainsi P inversible telle que PA est échelonnée réduite.

À présent, montrons que la réunion des orbites est disjointe, c'est-à-dire que deux matrices échelonnées en lignes réduites E et E' dans la même orbite sont égales.

Première preuve (élémentaire). Notons $(j_1, \dots, j_r)$ et $(j'_1, \dots, j'_r)$ les types de E et E' : ils ont pour longueur commune le rang de E et E'. Soit P une matrice de $\mathrm{GL}_m(\mathbb{K})$ telle que $PE = E'$. Les $j_1 - 1$ premières colonnes de PE sont nulles ; celle d'indice j_1 est la première colonne de P. Or, la première colonne non nulle de E' est ${}^t(1\ 0\ \cdots\ 0)$ et elle est en position j'_1. D'où l'égalité $j_1 = j'_1$ et l'écriture :

$$P = \left(\begin{array}{c|ccc} 1 & p_{12} & \cdots & p_{1n} \\ \hline 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{array}\right).$$

Découpons E et E' par blocs rectangulaires de la façon suivante :

$$E = \left(\frac{L}{F}\right), \quad E' = \left(\frac{L'}{F'}\right), \quad \text{où} \quad L, L' \in \mathscr{M}_{1,n}(\mathbb{K}), \ F, F' \in \mathscr{M}_{m-1,n}(\mathbb{K}).$$

L'égalité $PE = E'$ donne, en regardant le bloc du bas : $QF = F'$. Or, la définition des matrices échelonnées en lignes réduites fait que F et F' le sont ! Par une récurrence passée sous silence, on obtient : $F = F'$ et $Q = \mathrm{I}_{m-1}$.

Enfin, le coefficient d'indice $(1, j_k)$ de E' est p_{1k} d'une part et 0 d'autre part : on en déduit que $P = \mathrm{I}_m$ et que $E = E'$.

Deuxième preuve (un peu plus conceptuelle). Par l'assertion (i), on a : $\operatorname{Ker} E = \operatorname{Ker} E'$. On notera K cet espace et $K^\perp$ l'orthogonal de K dans le dual de $\mathbb{K}^n$: il s'identifie[3] à l'espace engendré par les lignes de E (ou celles de E'). Notons $(e_i^*)_{1 \leqslant i \leqslant m}$ la base canonique duale de $(\mathbb{K}^m)^*$.

Commençons par montrer que E et E' ont le même type. Étant donné une forme linéaire sur $\mathbb{K}^n$, disons $\ell = \sum_{j=1}^n y_j e_j^*$, on note $\text{min-supp}(\ell)$ le plus petit indice j tel que y_j n'est pas nul. On constate alors que puisque E et E' sont échelonnées, on peut repérer j_1 (resp. j'_1) comme le plus petit indice d'une combinaison linéaire des lignes de E (resp. E') :

$$j_1 = \min\{\text{min-supp}\,\ell \ , \ \ \ell \in K^\perp\} = j'_1.$$

Mais, E et E' sont de plus réduites, si bien que l'on peut repérer j_2 grâce aux éléments de $K^\perp$ dont le coefficient d'indice j_1 est nul :

$$j_2 = \min\{\text{min-supp}\,\ell \ , \ \ \ell \in K^\perp \cap (e_{j_1}^*)^\perp\} = j'_2.$$

[3] Si on pense $\mathbb{K}^n$ comme l'espace $\mathscr{M}_{n,1}(\mathbb{K})$ des colonnes, son dual $(\mathbb{K}^n)^*$ s'identifie à l'espace $\mathscr{M}_{1,n}(\mathbb{K})$ des lignes. On notera typiquement un élément du dual de $\mathbb{K}^m$ par la lettre ℓ, pour « ligne » ou « forme linéaire ». Le couplage entre le dual de $\mathbb{K}^n$ et $\mathbb{K}^n$ est l'application $\mathscr{M}_{1,n}(\mathbb{K}) \times \mathscr{M}_{n,1}(\mathbb{K}) \to \mathbb{K}$, $(\ell, v) \mapsto \ell v$. Le j^{e} vecteur de la base canonique duale e_j^* s'identifie à la ligne $(0\ \cdots\ 0\ 1\ 0 \cdots\ 0)$ (le coefficient 1 à la j^{e} position).

De proche en proche, pour k compris entre 1 et r :

$$j_k = \min\{\text{min-supp}\,\ell\ ,\ \ \ell \in K^\perp \cap (e^*_{j_1})^\perp \cap \cdots \cap (e^*_{j_{k-1}})^\perp\} = j'_k.$$

Ainsi, E et E' ont le même type $(j_1, \ldots, j_r)$.

Enfin, prouvons que E et E' sont égales. On associe aux lignes de la matrice $E = (a_{ij})$ les équations de l'orthogonal $K^\perp$ du noyau de E : pour i compris entre 1 et r, on définit la forme linéaire $\ell_i = \sum_{j=1}^n a_{ij} e^*_j$. On définit de même les ℓ'_i pour E'. Notons enfin J le complémentaire de $\{j_1, \ldots, j_r\}$ dans $\{1, \ldots, n\}$. Les familles $(\ell_i)_{1 \leqslant i \leqslant r}$ et $(\ell'_i)_{1 \leqslant i \leqslant r}$ sont deux bases de $K^\perp$. La matrice de passage de la première à la seconde s'écrit $(p_{ij})_{1 \leqslant i,j \leqslant r}$, de sorte que : $\ell'_j = \sum_i p_{ij} \ell_i$. En évaluant en $e_{j_1}, \ldots, e_{j_r}$, on obtient : $p_{ij} = \delta_{ij}$ (le δ de Kronecker) pour tout couple (i, j). □

2.4. Remarque. La deuxième fin de preuve consistait essentiellement à remarquer que le type d'une matrice (échelonnée) est caractérisé par la position relative du sous-espace engendré par les lignes (*i.e.* l'orthogonal du noyau) par rapport à un drapeau complet adapté à la base (en termes de dual, le k^{e} sous-espace est $\mathbb{K}e^*_1 \oplus \cdots \oplus \mathbb{K}e^*_k$). Voir ci-dessous une version en colonnes plus lisible.

Par transposition, on obtient le théorème dual, qui précise le lemme 1.2. En effet, il suffit de remarquer que par transposition, l'orthogonal du noyau, c'est-à-dire l'espace engendré par les lignes, devient l'espace engendré par les colonnes, c'est-à-dire l'image de la matrice. On donne cependant une preuve un peu abrégée pour éviter la gymnastique de lire la preuve avec la tête penchée à 90° dans un miroir.

2.5. Théorème. *Soit n et m deux entiers. On considère l'action de $\mathrm{GL}_m(\mathbb{K})$ par multiplication à droite sur l'espace $\mathscr{M}_{n,m}(\mathbb{K})$ des matrices $n \times m$. Alors :*

(i) *deux matrices A et A' de $\mathscr{M}_{n,m}(\mathbb{K})$ sont dans la même orbite sous $\mathrm{GL}_m(\mathbb{K})$ si et seulement si elles ont la même image :*

$$\exists Q \in \mathrm{GL}_m(\mathbb{K}),\ A' = AQ \iff \mathrm{Im}\, A = \mathrm{Im}\, A';$$

(ii) *toute orbite contient une unique matrice échelonnée réduite en colonnes :*

$$\mathscr{M}_{n,m}(\mathbb{K}) = \bigcup_{E \in \mathscr{E}^\vee_{n,m}} E \cdot \mathrm{GL}_m(\mathbb{K}),$$

où la réunion est disjointe et où $\mathscr{E}^\vee_{n,m}$ désigne l'ensemble des matrices échelonnées en colonnes réduites de taille $n \times m$.

Démonstration. (i) Si $A' = AQ$ pour Q inversible, les colonnes de A' sont des combinaisons linéaires de celles de A donc l'image de A' est incluse dans celle de A. L'inclusion inverse vient de l'égalité $A = A'Q^{-1}$.

Réciproquement, supposons que les images de A et A' sont égales. Soit φ_A et $\varphi_{A'} : \mathbb{K}^m \to \mathbb{K}^n$ les multiplications par A et A'. Soit $(f_1, \dots, f_r)$ une base de l'image commune (où $r = \operatorname{rg} A = \operatorname{rg} A'$). On choisit des antécédents $(e_1, \dots, e_r)$ (resp. $(e'_1, \dots, e'_r)$) de $(f_1, \dots, f_r)$ par A (resp. par A') ; on choisit un endomorphisme ψ de $\mathbb{K}^m$ qui, pour k compris entre 1 et r, envoie e'_k sur e_k et qui envoie bijectivement $\operatorname{Ker} \varphi_{A'}$ sur $\operatorname{Ker} \varphi_A$. Par construction, on a l'égalité $\varphi_{A'} = \varphi_A \psi$ qui devient, dans les bases canoniques : $A' = AQ$ avec Q inversible.

(ii) La stratégie de l'algorithme du pivot de Gauss permet, par des manipulations sur les colonnes d'une matrice A, de la transformer en une matrice échelonnée puis en une matrice échelonnée réduite E. Cela implique l'existence de Q inversible telle que $AQ = E$. Ainsi, toute orbite contient une matrice échelonnée en colonnes réduite.

Supposons que E et E' soient deux matrices échelonnées en colonnes réduites dans la même orbite sous $\mathrm{GL}_m(\mathbb{K})$. Par le point (i), ces deux matrices ont la même image, notée I. Montrons qu'elles ont le même type. Soit $i_1, \dots, i_r$ les indices des lignes des pivots de E, idem pour E'. Pour j entier compris entre 0 et m, soit[4] $V_j^\vee$ l'espace engendré par les vecteurs $e_{j+1}, \dots, e_n$ de la base canonique de $\mathbb{K}^n$ (c'est l'intersection des noyaux de $e_1^*, \dots, e_j^*$; $V_0^\vee = \mathbb{K}^n$). On caractérise i_1 comme le plus petit indice d'un coefficient non nul d'un élément de $\operatorname{Im}(E)$, idem pour i'_1 :

$$i_1 = \max\left\{ j \in \{1, \dots, n\} \ , \ \ I \subset V_j^\vee \right\} = i'_1.$$

Plus généralement, on montre par récurrence finie que pour k compris entre 2 et r :

$$i_k = \max\left\{ j \in \{1, \dots, n\} \ , \ \ I \cap V_{i_{k-1}}^\vee \subset V_j^\vee \right\} = i'_k.$$

Ainsi, E et E' ont le même type.

Soit Q est une matrice inversible telle que $E' = EQ$. Plissons les yeux et ne regardons que les lignes d'indices $i_1, \dots, i_r$: la sous-matrice correspondante de E et E' est la juxtaposition de l'identité I_r et de la matrice nulle de taille $(m-r)\times r$; celle de EQ est donc formée des r premières lignes de Q, et vu

[4] Le $\vee$ de la notation $V_i^\vee$ sera justifié plus loin par l'introduction d'autres sous-espaces V_i qui sont, pour le cœur de ce chapitre, plus importants.

que l'on a $E' = EQ$, on peut écrire Q sous la forme $(EQ = E')$:

$$Q = \begin{pmatrix} \mathrm{I}_r & 0 \\ * & * \end{pmatrix}.$$

Il en résulte que $E' = EQ = E$. □

3. Décomposition en cellules des grassmanniennes

Ce paragraphe aboutit à une sorte de reformulation géométrique du théorème 2.5. Dans le paragraphe précédent, l'accent était mis sur les matrices ; ici, on ne s'intéresse qu'à l'espace engendré par les colonnes et l'on se restreint aux matrices de rang maximal.

3.1. Deux présentations de la grassmannienne

Remarquons que l'action du groupe $\mathrm{GL}_n(\mathbb{K})$ par multiplication à gauche sur $\mathscr{M}_{n,m}(\mathbb{K})$ commute avec l'action à droite de $\mathrm{GL}_m(\mathbb{K})$: c'est l'associativité —au sens large— du produit matriciel. Par suite, le quotient $\mathscr{M}^0_{n,m}(\mathbb{K})/\mathrm{GL}_m(\mathbb{K})$ hérite d'une action à gauche de $\mathrm{GL}_n(\mathbb{K})$; de plus, l'application $\overline{\varphi}$ commute aux actions de $\mathrm{GL}_n(\mathbb{K})$: pour P matrice inversible $n\times n$ et A matrice $n\times m$ de rang maximal, on a :

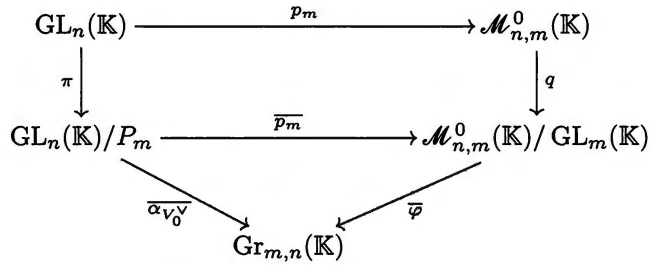
$$P\big(\overline{\varphi}(A)\big) = \overline{\varphi}(PA).$$

Cette formule hermétique traduit simplement que, si l'on désigne par f_j $(1 \leqslant j \leqslant m)$ les colonnes de A, on a :

$$P \cdot (\mathbb{K}f_1 \oplus \cdots \oplus \mathbb{K}f_m) = \mathbb{K}Pf_1 \oplus \cdots \oplus \mathbb{K}Pf_m.$$

Pour les considérations topologiques ultérieures, on va vérifier que l'application $\overline{\varphi}$, qui établit une bijection de $\mathscr{M}^0_{m,n}(\mathbb{K})/\mathrm{GL}_m(\mathbb{K})$ sur $\mathrm{Gr}_{m,n}(\mathbb{K})$ est aussi un homéomorphisme. Pour cela, il faut se rappeler, du paragraphe II-4.3, que la topologie de la grassmannienne vient de sa réalisation comme quotient $\mathrm{GL}_n(\mathbb{K})/P_m$, où P_m est le stabilisateur du sous-espace engendré par les m premiers vecteurs de la base canonique.

L'ingrédient « nouveau » est la projection $p_m : \mathrm{GL}_n(\mathbb{K}) \to \mathscr{M}^0_{m,n}$ qui, à un élément g, associe la sous-matrice de g formée des m premières colonnes (voir la remarque II-4.3.3). Vu la forme de P_m, un calcul matriciel par blocs permet de vérifier que la composée de p_m et de la projection canonique $q : \mathscr{M}^0_{n,m}(\mathbb{K}) \to \mathscr{M}^0_{n,m}(\mathbb{K})/\mathrm{GL}_m(\mathbb{K})$ passe au quotient par P_m, ce qui donne le diagramme commutatif suivant :



Lorsque $\mathbb{K}$ est le corps des réels ou des complexes, p_m est continue et ouverte. Mais c'est aussi le cas des projections canoniques π et q. Par les propriétés des topologies quotients, il en résulte que $\overline{p_m}$ est continue et ouverte. Comme on sait qu'elle est bijective, c'est un homéomorphisme. Ainsi, $\overline{\varphi}$ en est un aussi.

3.1.1. Remarque. Pour $m = 1$, l'espace $\mathscr{M}^0_{n,1}(\mathbb{K})$ n'est autre que $\mathbb{K}^n \setminus \{0\}$ et $\mathrm{GL}_1(\mathbb{K})$ est $\mathbb{K}^*$. Ce qui précède étend aux grassmanniennes la présentation des espaces projectifs vue à l'exercice II-F.23.

3.2. Première version de la décomposition : par les matrices co-échelonnées

On a utilisé la définition standard des matrices échelonnées en colonnes. Cependant, pour l'étude géométrique qui vient, elle n'est pas très adaptée et donne une « mauvaise » définition de type (et de mauvaises formules de dimension). C'est que le type provient de la position de l'image de la matrice par rapport au drapeau $(V_i^\vee)_{0 \leqslant i \leqslant n}$ qui n'est pas adapté à la base canonique $(e_1, \ldots, e_n)$, mais à la base $(e_n, \ldots, e_1)$. On rectifie en conséquence les choses — moralement, cela revient à faire agir l'involution $(1, n)(2, n-1)\cdots$.

3.2.1. Définition. Dans ce livre, on convient de dire qu'une matrice est *co-échelonnée* si elle vérifie les conditions ci-après :
- si une colonne est nulle, toutes les colonnes suivantes sont nulles ;
- le pivot d'une colonne est strictement plus bas que les pivots des lignes précédentes.

Une matrice co-échelonnée est dite *réduite* si, de plus,
- tous les pivots sont égaux à 1 et sont les seuls coefficients non nuls de leur ligne.

On appelle *type* d'une matrice co-échelonnée la liste strictement croissante $\mathbf{i} = (i_1, i_2, \ldots, i_m)$ des indices des lignes des pivots. (Il revient au même de parler de la partie $\{i_1, \ldots, i_m\}$.) On note $\mathscr{E}^\wedge_{\mathbf{i}}$ l'ensemble des matrices co-échelonnées réduites de type $\mathbf{i}$ (comparer à $\mathscr{E}^\vee_{n,m}$ ci-dessus).

3.2.2. Proposition. *Soit n et m deux entiers avec $m \leqslant n$. Soit F un sous-espace de $\mathbb{K}^n$ de dimension m. Il existe une unique suite strictement croissante $\mathbf{i} = (i_1, \ldots, i_m)$ à valeurs dans $\{1, \ldots, n\}$ et une unique matrice co-échelonnée réduite de type $\mathbf{i}$ dont les colonnes engendrent F.*

Comme c'est la troisième démonstration de ce résultat, on va être un peu plus expéditif... Schématiquement, par des opérations sur les colonnes, on fait :

$$\begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{pmatrix} \rightsquigarrow \begin{pmatrix} * & * & * \\ * & \neq 0 & * \\ \neq 0 & 0 & * \\ 0 & 0 & * \\ 0 & 0 & \neq 0 \\ 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} * & * & * \\ * & 1 & * \\ 1 & 0 & * \\ 0 & 0 & * \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \rightsquigarrow$$

$$\rightsquigarrow \begin{pmatrix} * & * & * \\ 1 & * & * \\ 0 & 1 & * \\ 0 & 0 & * \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} * & * & * \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & * \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Démonstration. Première étape (existence). Soit A une matrice $n \times m$ dont les colonnes engendrent F. Par des opérations sur les colonnes de A, on transforme A en une matrice A' où le nombre total de zéros apparaissant sous les pivots est maximal. Les lignes où se trouvent les pivots sont toutes distinctes : en effet, si deux pivots de deux colonnes étaient sur la même ligne, une combinaison linéaire adéquate permettrait d'annuler le pivot de l'une d'entre elles, ce qui contredit la maximalité du nombre de zéros de A'.

En multipliant les colonnes de A' par des coefficients convenables, on fait en sorte que les pivots vaillent tous 1 ; en permutant certaines colonnes, on s'arrange pour que les indices des lignes des pivots croissent de la gauche à la droite ; par des combinaisons linéaires convenables, on annule tous les coefficients situés à droite d'un pivot (sur la même ligne). La matrice A'' obtenue est co-échelonnée réduite et ses colonnes engendrent F, comme les colonnes de A.

Deuxième étape (unicité du type). Pour i compris entre 0 et n, soit V_i le sous-espace engendré par les i premiers vecteurs de la base canonique de $\mathbb{K}^n$. Considérons la suite de dimensions :

$$0 = \dim(F \cap V_0) \leqslant \dim(F \cap V_1) \leqslant \cdots \leqslant \dim(F \cap V_{n-1}) \leqslant \dim(F \cap V_n) = m.$$

Comme $F \cap V_i/F \cap V_{i-1}$ s'injecte[5] dans la droite V_i/V_{i-1}, les sauts de dimensions[6] valent au plus 1. Il y en a donc exactement m, qui se produisent pour les valeurs $i_1 < \cdots < i_m$ de i.

Or, pour calculer ces intersections, on peut utiliser la matrice A''. On voit que $\dim F \cap V_i$ est le nombre de colonnes dont le pivot est sur une ligne d'indice inférieur ou égal à i ; autrement dit, les sauts de dimension, qui ne dépendent que de F, se produisent exactement aux indices des lignes des pivots. Cela prouve l'unicité de $\mathbf{i} = (i_1, \ldots, i_m)$. □

3.2.3. Définition. Avec les notations de la proposition, on dit que F est un *sous-espace de type* $\mathbf{i}$. On note $\mathscr{O}_\mathbf{i}$ l'ensemble des sous-espaces de type $\mathbf{i}$.

3.2.4. Théorème (Décomposition en cellules de la grassmannienne). *Soit n et m deux entiers avec $m \leqslant n$.*

(i) *La grassmannienne $\mathrm{Gr}_{m,n}(\mathbb{K})$ est la réunion disjointe des $\mathscr{O}_\mathbf{i}$, où $\mathbf{i}$ parcourt les suites strictement croissantes $1 \leqslant i_1 < \cdots < i_m \leqslant n$ (*i.e. *les parties à m éléments de $\{1, \ldots, n\}$).*

(ii) *Pour un type $\mathbf{i} = (i_1, \cdots, i_m)$ fixé, il existe une bijection, qui est un homéomorphisme lorsque cela a un sens, $\mathbb{K}^{|\mathbf{i}|} \to \mathscr{O}_\mathbf{i}$, où*

$$|\mathbf{i}| = \sum_{j=1}^{m} (i_j - j).$$

Démonstration. L'assertion (i) est une reformulation de la proposition.

L'application $\mathscr{E}_\mathbf{i}^\wedge \to \mathscr{O}_\mathbf{i}$ est bijective ; c'est la restriction de l'application $\mathscr{M}_{n,m}^0 \to \mathrm{Gr}_{m,n}(\mathbb{K})$, qui est continue et ouverte (théorème d'homéomorphisme II-3.4.3). Pour terminer la preuve du théorème, il suffit donc de vérifier que $\mathscr{E}_\mathbf{i}^\wedge$ est un espace affine de dimension $|\mathbf{i}|$. Pour cela, il suffit de compter le nombre de coefficients non concernés par les conditions qui définissent les matrices co-échelonnées : sur la colonne d'indice j, il y a $i_j - 1$ coefficients au-dessus du pivot, dont $j - 1$ valent 0 à cause des pivots des colonnes précédentes ; restent donc $i_j - j$ degrés de liberté. □

Par abus de notation, on résumera les résultats de ce théorème en la formule suivante :

$$\mathrm{Gr}_{m,n}(\mathbb{K}) = \bigsqcup_\mathbf{i} \mathbb{K}^{|\mathbf{i}|},$$

où $\mathbf{i}$ parcourt les parties à m éléments de $\{1, \ldots, n\}$.

[5] La composée $F \cap V_i \to V_i \to V_i/V_{i-1}$ admet pour noyau $F \cap V_{i-1}$, et induit donc l'injection mentionnée.

[6] C'est-à-dire les différences non nulles $\dim(F \cap V_i) - \dim(F \cap V_{i-1})$.

On retrouvera au second tome, dans le chapitre IV-1, une application de ce résultat à de jolies formules combinatoires.

Lorsque l'on décompose un espace topologique en une réunion disjointes de parties C_i, où chaque C_i est homéomorphe à un espace affine, on appellera *cellule* chaque C_i et la décomposition sera appelée *décomposition en cellules.*

3.3. Deuxième version de la décomposition : cellules comme B-orbites

Considérons le problème d'algèbre linéaire élémentaire suivant : étant donné une matrice A de taille $n \times m$ et de rang maximal m, trouver une forme normale lorsque l'on autorise les manipulations sur les rangées suivantes :
- produit d'une ligne par une constante,
- combinaison linéaire de la forme $\ell_i \to \ell_i + \alpha \ell_j$, $i < j$, c'est-à-dire que l'on remplace une ligne par une combinaison linéaire utilisant une ligne située *en-dessous.*
- toutes les opérations habituelles sur les colonnes.

Par rapport à la situation envisagée dans le chapitre I pour le théorème du rang, on restreint les opérations sur les lignes ; par rapport à la situation des grassmanniennes, on ajoute certaines opérations sur les lignes.

Traduisons cela en termes de groupes. Notons B le sous-groupe de $\mathrm{GL}_n(\mathbb{K})$ formé des matrices triangulaires supérieures[7] :

$$B = \begin{pmatrix} * & \cdots & \cdots & * \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & * \end{pmatrix} \subset \mathrm{GL}_n(\mathbb{K}).$$

On dispose sur $\mathscr{M}^0_{n,m}(\mathbb{K})$ de deux actions qui commutent entre elles :

$$B \curvearrowright \mathscr{M}^0_{n,m}(\mathbb{K}) \curvearrowleft \mathrm{GL}_m(\mathbb{K}),$$

une du groupe B par multiplication à gauche, l'autre de $\mathrm{GL}_m(\mathbb{K})$ par multiplication à droite. On en déduit une action du groupe produit définie, pour $(b, g) \in B \times \mathrm{GL}_m(\mathbb{K})$ et $A \in \mathscr{M}^0_{n,m}(\mathbb{K})$ par : $(b, g) \cdot A = bAg^{-1}$.

Il est d'usage de noter le quotient ainsi :

$$B \backslash \mathscr{M}^0_{n,m}(\mathbb{K}) / \mathrm{GL}_m(\mathbb{K}) = \mathscr{M}^0_{n,m}(\mathbb{K}) / \big(B \times \mathrm{GL}_m(\mathbb{K})\big).$$

Par l'exercice B.2, plutôt que l'action du produit, il revient au même d'étudier l'action de B sur $\mathscr{M}^0_{n,m}(\mathbb{K}) / \mathrm{GL}_m(\mathbb{K})$ ou, en identifiant ce quotient avec

[7] La lettre B est l'initiale du nom d'Armand Borel, qui a repéré l'importance de ce sous-groupe et en a défini des analogues pour les autres groupes classiques.

la grassmannienne $\mathrm{Gr}_{m,n}(\mathbb{K})$, puis avec $\mathrm{GL}_n(\mathbb{K})/P_m$, nous sommes en train d'étudier *les B-orbites de* $\mathrm{GL}_n(\mathbb{K})/P_m$.

3.3.1. Proposition. *Soit n et m deux entiers avec $m \leqslant n$. Soit P_m le stabilisateur de l'espace engendré par les m premiers vecteurs de la base canonique de $\mathbb{K}^n$ et B le stabilisateur du drapeau standard de $\mathbb{K}^n$. Alors les B-orbites dans $\mathrm{GL}_n(\mathbb{K})/P_m$ s'identifient aux cellules de la grassmannienne $\mathrm{Gr}_{m,n}(\mathbb{K})$.*

Démonstration. Soit $\mathscr{O}_{\mathbf{i}}$ la cellule de type $\mathbf{i} = (i_1, \ldots, i_m)$. Soit F un élément de $\mathscr{O}_{\mathbf{i}}$ et A une matrice échelonnée réduite dont les colonnes engendrent F. Alors, A est dans la B-orbite de la matrice échelonnée réduite $E_{\mathbf{i}}$ de type $\mathbf{i}$ dont tous les coefficients autres que les pivots sont nuls (version élémentaire : on transforme $E_{\mathbf{i}}$ en A en faisant des combinaisons linéaires de lignes situées en-dessous). Donc, une cellule est contenue dans une B-orbite.

Inversement, si F et F' sont dans la même B-orbite, c'est-à-dire si $F' = bF$ pour un élément b de B, alors leur type est le même : en effet, puisque b fixe le drapeau de référence, on a :

$$\forall k \in \{1, \ldots, m\}, \quad \dim(F \cap V_k) = \dim(bF \cap bV_k) = \dim(F' \cap V_k),$$

et on sait que le type est déterminé par les sauts de cette suite de dimensions. Ainsi, une B-orbite est contenue dans une cellule de la grassmannienne. D'où la proposition. □

3.3.2. Remarque. Les matrices échelonnées habituelles, c'est-à-dire les éléments de $\mathscr{E}_{\mathbf{i}}^{\vee}$, correspondent à la décomposition de la grassmannienne en orbites sous le groupe B^- des matrices triangulaires inférieures. En rapport avec la décomposition de Bruhat (voir le second tome), on peut remarquer que $B^- = \sigma_0 B \sigma_0$, où σ_0 est l'unique élément ayant $n(n-1)/2$ inversions.

3.4. Adhérence des cellules : position du problème

À présent, on suppose que $\mathbb{K}$ est le corps des réels ou des complexes. On s'intéresse à l'adhérence des cellules des grassmanniennes. D'une part, c'est joli. D'autre part, pour déduire des résultats globaux sur la topologie de la grassmannienne à partir d'une partition, il faut vérifier que les parties sont bien agencées les unes par rapport aux autres. On verra une troisième motivation à la fin du chapitre.

Dans toute action continue d'un groupe sur un espace topologique, l'adhérence d'une partie stable est stable (exercice II-F.4). En particulier, l'adhérence d'une orbite est une réunion d'orbites. Ici, comme les cellules des

grassmanniennes sont aussi des B-orbites, on sait déjà que l'adhérence d'une cellule est une réunion de cellules.

Nous allons constater que la relation « appartenir à l'adhérence de », *i.e.* l'ordre de dégénérescence, est une relation d'ordre sur les cellules. Ce n'est pas évident *a priori*, même pour des orbites pour l'action d'un groupe : les parties $\mathbb{Q}$ et $\mathbb{R} \setminus \mathbb{Q}$, qui sont pourtant des réunions d'orbites sous l'action du groupe additif $\mathbb{Q}$ par translation, sont dans l'adhérence l'une de l'autre ; c'est que l'action de $\mathbb{Q}$ n'est pas une action propre.

3.5. Adhérence des cellules : le cas de $\mathbf{Gr_{2,4}}$

On s'intéresse aux plans de $\mathbb{K}^4$, c'est-à-dire aux droites de l'espace projectif $\mathbb{P}^3(\mathbb{K})$. Le drapeau de référence est

$$\{0\} = V_0 \subset V_1 \subset V_2 \subset V_3 \subset V_4 = \mathbb{K}^4, \quad \text{où } V_i = \mathbb{K}e_1 \oplus \cdots \oplus \mathbb{K}e_i \ (0 \leqslant i \leqslant 4).$$

Il y a $\binom{4}{2}$ = six types possibles. Voici les matrices échelonnées et la caractérisation des sous-espaces correspondants, disposées selon l'ordre de dégénérescence que nous verrons juste après.

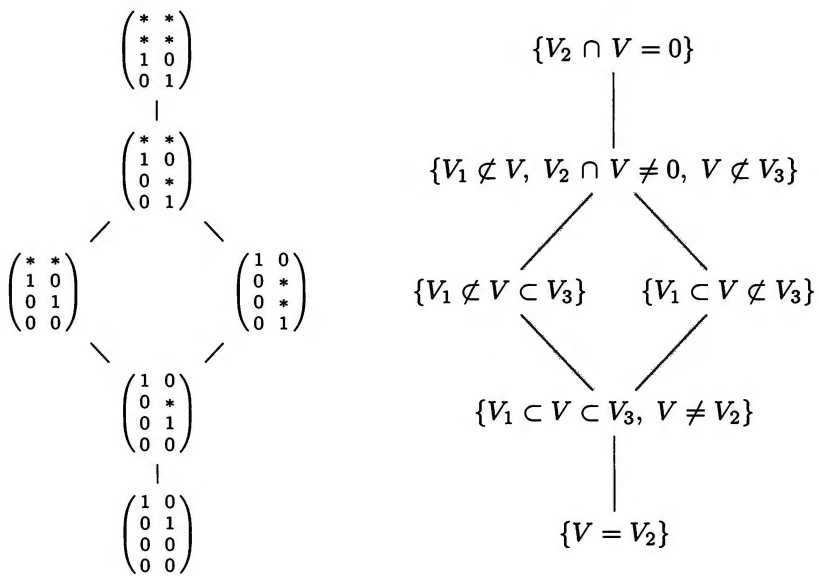


Figure 3.1. *Adhérence des cellules dans* $\mathrm{Gr}_{4,2}$

Il est facile de voir que si $i_1 \leqslant j_1$ et $i_2 \leqslant j_2$, la cellule $\mathscr{O}_{i_1,i_2}$ est dans l'adhérence de la cellule $\mathscr{O}_{j_1,j_2}$. Par exemple, pour $(i_1,i_2) = (1,3)$ et $(j_1,j_2) = (2,3)$,

on considère :

$$\forall \varepsilon \in \mathbb{K}, \quad A(\varepsilon) = \begin{pmatrix} 1 & 0 \\ \varepsilon & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Pour $\varepsilon \neq 0$, $A(\varepsilon)$ appartient à $\mathscr{O}_{2,3}$ et, lorsque ε tend vers 0, $A(\varepsilon)$ tend vers un élément de $\mathscr{O}_{1,3}$.

En revanche, les cellules $\mathscr{O}_{2,3}$ et $\mathscr{O}_{1,4}$ ne sont pas dans l'adhérence l'une de l'autre. Si (A_k) est une suite convergente d'éléments de $\mathscr{O}_{2,3}$, le mineur correspondant aux lignes d'indices $(1,4)$ de A_k est nul (la dernière ligne de A_k est nulle) donc la limite n'appartient pas à $\mathscr{O}_{1,4}$: en effet, les mineurs d'indices $(1,4)$ des éléments de $\mathscr{O}_{1,4}$ ne sont pas nuls. On peut échanger les rôles de $(1,4)$ et $(2,3)$ ici.

3.5.1. Remarque. Cet exemple a un intérêt historique : c'est celui qui a motivé Schubert à introduire les *variétés de Schubert*, qui sont les adhérences des cellules $\mathscr{O}_{i_1,i_2}$. Sa motivation était la *géométrie énumérative* dans l'esprit de l'énoncé suivant : il y a exactement deux droites de $\mathbb{P}^3$ qui s'appuient sur quatre droites données en position générale. Voir [30].

3.6. Adhérence des cellules : le cas général

3.6.1. Proposition. *Soit n et m deux entiers avec $m \leqslant n$. Étant donné deux suites strictement croissantes $\mathbf{i}' = (i'_1, \ldots, i'_m)$ et $\mathbf{i} = (i_1, \ldots, i_m)$ à valeurs dans $\{1, \ldots, n\}$, il est équivalent de dire :*

(i) $\mathscr{O}_{\mathbf{i}'} \subset \overline{\mathscr{O}_{\mathbf{i}}}$;

(ii) $\forall k \in \{1, \ldots, m\}$, $i'_k \leqslant i_k$.

En particulier, la relation « être inclus dans l'adhérence de » est une relation d'ordre sur les cellules des grassmanniennes.

Démonstration. Pour A matrice $n \times m$, on note $A_{\mathbf{i}'}$ la sous-matrice de A formée des lignes $i'_1, \ldots, i'_m$ et $\Delta_{\mathbf{i}'}(A) = \det A_{\mathbf{i}'}$ le mineur correspondant. Pour tout élément F de la grassmannienne, choisissons une matrice A dont les colonnes engendrent F. Le fait que $\Delta_{\mathbf{i}'}(A)$ soit nul ou pas ne dépend que de F et pas de A et définit un fermé de la grassmannienne. En effet, si A' est une autre matrice, on a l'égalité $A' = Ah$ pour h de $\mathrm{GL}_m(\mathbb{K})$, ce qui donne $A'_{\mathbf{i}'} = A_{\mathbf{i}'}h$. Ainsi, la partie $\Delta_{\mathbf{i}'}^{-1}(0)$ de $\mathscr{M}_{n,m}(\mathbb{K})$ est stable par $\mathrm{GL}_m(\mathbb{K})$ et, par continuité de $\Delta_{\mathbf{i}'}$ c'est un fermé.

Supposons qu'il existe k tel que $i'_k > i_k$. Lorsque F appartient à $\mathscr{O}_{\mathbf{i}}$, la ligne d'indice i'_k de A est nulle, si bien que $\Delta_{\mathbf{i}'}(A) = 0$. En revanche, si F appartient à $\mathscr{O}_{\mathbf{i}'}$, on a : $\Delta_{\mathbf{i}'}(A) \neq 0$ (si, par exemple, A est la matrice échelonnée réduite de $\mathscr{O}_{\mathbf{i}'}$, on a : $\Delta_{\mathbf{i}'}(A) = 1$). Par conséquent, aucun élément de $\mathscr{O}_{\mathbf{i}'}$ ne peut être dans l'adhérence de $\mathscr{O}_{\mathbf{i}}$.

Réciproquement, on suppose que $i'_k \leqslant i_k$ pour tout k. Pour ε scalaire, considérons la matrice $A(\varepsilon)$ dont la k^{e} colonne est, pour k compris entre 1 et m :

$$ {}^t(0 \quad \cdots \quad 0 \underset{(i'_k)}{1} 0 \quad \cdots \quad 0 \underset{(i_k)}{\varepsilon} 0 \quad \cdots \quad 0). $$

Lorsque ε n'est pas nul, $A(\varepsilon)$ appartient à $\mathscr{O}_{\mathbf{i}}$. Lorsque ε tend vers zéro, $A(\varepsilon)$ tend vers une matrice échelonnée réduite de $\mathscr{O}_{\mathbf{i}'}$. Il en résulte que l'adhérence de $\mathscr{O}_{\mathbf{i}}$ contient un élément de $\mathscr{O}_{\mathbf{i}'}$ puis, vu que l'adhérence d'une B-orbite est une réunion de B-orbites (lemme II-3.1.4), que $\mathscr{O}_{\mathbf{i}'}$ est incluse dans l'adhérence de $\mathscr{O}_{\mathbf{i}}$.

La dernière affirmation provient de ce que les cellules sont paramétrées par leur type et que la relation de l'assertion (ii) est une relation d'ordre (ordre lexicographique sur les m-listes). □

3.7. Caractérisation géométrique des cellules et leurs adhérences

3.7.1. Corollaire. *On garde les notations de la proposition* 3.6.1. *Soit* $\mathbf{i} = (i_1, \ldots, i_m)$ *un type. Alors :*

$$ \mathscr{O}_{\mathbf{i}} = \{F \in \mathrm{Gr}_{n,m}(\mathbb{K}) \ , \ \ \forall k \in \{1, \ldots, k\}, \ \dim F \cap V_k = i_k\}, $$
$$ \overline{\mathscr{O}_{\mathbf{i}}} = \{F \in \mathrm{Gr}_{n,m}(\mathbb{K}) \ , \ \ \forall k \in \{1, \ldots, k\}, \ \dim F \cap V_k \geqslant i_k\}. $$

A. Annexe. Opérations élémentaires sur les rangées

Soit p un entier naturel. On note I_p la matrice identité d'ordre p. Pour i et j entiers compris entre 1 et p, on note $E_{i,j}^{(p)}$ la matrice $p \times p$ dont le coefficient d'indice (i, j) vaut 1 et les autres 0. On considère les matrices $p \times p$ suivantes

(D pour dilatation, T pour transvection, P pour permutation) :

$$1 \leqslant i \leqslant p,\ \alpha \in \mathbb{K},\ \alpha \neq 0 \qquad D_{i,\alpha}^{(p)} = \begin{pmatrix} \mathrm{I}_{i-1} & & \\ & \alpha & \\ & & \mathrm{I}_{p-i-1} \end{pmatrix},$$

$$1 \leqslant i \neq j \leqslant p,\ \beta \in \mathbb{K},\ \beta \text{ en position } (i,j) \qquad T_{i,j;\beta}^{(p)} = \mathrm{I}_p + \beta E_{i,j}^{(p)} = \begin{pmatrix} 1 & & & \\ & \ddots & \beta & \\ & & \ddots & \\ & & & 1 \end{pmatrix},$$

$$1 \leqslant i < j \leqslant p \qquad P_{i,j}^{(p)} = P_{j,i}^{(p)} = \begin{pmatrix} \mathrm{I}_{i-1} & & & & \\ & 0 & \dots & 1 & \\ & & \mathrm{I}_{j-i-1} & & \\ & 1 & \dots & 0 & \\ & & & & \mathrm{I}_{p-j-1} \end{pmatrix}.$$

On se donne aussi une matrice A de taille $m \times n$ à coefficients dans $\mathbb{K}$.

En calculant le produit $D_{i,\alpha}^{(m)} A$, on vérifie que l'on l'obtient en multipliant la i-ème ligne de A par α, ce que l'on notera : $L_i \leftarrow \alpha L_i$. Cela donne la première colonne du tableau suivant, dont on vérifiera les autres :

opération	$D_{i,\alpha}^{(m)} A$	$T_{i,j;\beta}^{(m)} A$	$P_{i,j}^{(m)} A$
résultat	$L_i \leftarrow \alpha L_i$	$L_i \leftarrow L_i + \beta L_j$	$L_i \leftrightarrow L_j$
opération	$A D_{i,\alpha}^{(n)}$	$A T_{i,j;\beta}^{(n)}$	$A P_{i,j}^{(n)}$
résultat	$C_i \leftarrow \alpha C_i$	$C_j \leftarrow C_j + \beta C_i$	$C_i \leftrightarrow C_j$

B. Exercices

B.1. Exercice (algorithme du pivot de Gauss et engendrement)

On veut montrer que le groupe spécial linéaire est engendré par les transvections. On reprend les notations du paragraphe A, *mais l'on se restreint aux opérations sur les rangées implémentées par les matrices de transvection. On fixe un entier n et un corps $\mathbb{K}$.*

1. *On commence par chercher des substituts aux permutations.*

   (a) *Vérifier la relation :*

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

(b) En déduire que, quel que soit l'emplacement du bloc 2×2, la matrice diagonale suivante est un produit de transvections :

$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & 0 & 1 & & \\ & & & -1 & 0 & & \\ & & & & & 1 & \\ & & & & & & \ddots \\ & & & & & & & 1 \end{pmatrix}.$$

2. On cherche des substituts aux dilatations. Soit λ un scalaire non nul.

(a) Montrer que la matrice

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}.$$

est un produit de transvections.

(b) En déduire que quel que soit l'emplacement de λ, la matrice diagonale suivante est un produit de transvections (le coefficient λ^{-1} est en dernière position) :

$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \\ & & & & & & & \lambda^{-1} \end{pmatrix}.$$

3. À l'aide de l'algorithme de Gauss, montrer que les transvections engendrent $\mathrm{SL}_n(\mathbb{K})$.

L'idée est d'appliquer les étapes habituelles de l'algorithme de Gauss avec les matrices ci-dessus, qui se substituent à celles qui déterminent les opérations élémentaires. On peut ainsi permuter deux rangées (quitte à introduire un signe), multiplier une rangée par une constante (pour peu que ce ne soit pas la dernière, quitte à modifier celle-ci) et faire des combinaisons linéaires de rangées. La seule chose à vérifier, c'est que l'on n'a pas besoin de faire une dilatation en arrivant à la dernière étape.

B.2. Exercice (Actions de deux groupes qui commutent). *Soit B et H deux groupes (topologiques) agissant (continûment) sur un ensemble (un espace topologique) X et soit $G = B \times H$. On suppose que les actions de B et H commutent, c'est-à-dire que pour $(b, c) \in B \times H$ et $x \in x$, on a : $b \cdot (c \cdot x) = c \cdot (b \cdot x)$.*

1. *Définir une action naturelle de G sur X et une action naturelle de B sur X/H.*
2. *On note $q : X \to X/H$ la projection canonique. Montrer que l'action de B commute à q, au sens où l'on a : $b \cdot q(x) = q(b \cdot x)$ pour tout b de B et x de X.*
3. *En déduire qu'il existe une bijection naturelle entre le quotient de X par G et le quotient de X/H par B.*

   Montrer que l'application qui, à la B-orbite de $q(x)$ dans X/H, associe la G-orbite de x dans X/G, est bien définie et bijective.

B.3. Exercice (Le groupe B comme sous-groupe résoluble maximal)

On rappelle qu'un groupe B est résoluble *s'il existe une famille finie de sous-groupes emboîtés*

$$B = B_0 \supset B_1 \supset \cdots \supset B_\ell = \{1\}$$

telle que pour tout entier k compris entre 0 et $\ell - 1$, le sous-groupe B_{k+1} soit distingué dans B_k et le quotient B_k/B_{k+1} soit abélien.

On fixe un entier naturel non nul n et l'on note B le groupe des matrices triangulaires supérieures inversibles.

1. *Montrer que le groupe dérivé de B, c'est-à-dire le groupe engendré par les commutateurs $[g, h] = ghg^{-1}h^{-1}$ $(g, h \in B)$, est le groupe U des matrices triangulaires unipotentes (dont toutes les valeurs propres sont égales à 1).*
2. *Montrer que B est résoluble.*

   On pourra prendre $B_1 = U$.
3. *Soit P un sous-groupe contenant strictement B. Montrer que P n'est pas résoluble.*

   On pourra commencer par le cas $n = 2$. Par des opérations élémentaires sur les rangées, montrer que P contient la matrice de la transposition (12). En déduire que $P = \mathrm{GL}_2(\mathbb{K})$. Adapter l'argument pour montrer qu'en général, P contient un sous-groupe isomorphe à $\mathrm{GL}_2(\mathbb{K})$.

B.4. Exercice (Décomposition LU)

*On rappelle que la «décomposition LU» consiste à essayer d'écrire une matrice inversible sous forme du produit d'une matrice triangulaire inférieure (*lower triangular *pour le L) et d'une matrice triangulaire supérieure (*upper triangular *pour le U). On reprend les notations B et U de l'exercice* B.3.

1. *Soit w_0 la (matrice de) permutation qui envoie* 1 *sur n,* 2 *sur $n-1$, etc. Vérifier que w_0 est une involution. Décrire le groupe $L = w_0 U w_0$.*
2. *Soit g un élément de $\mathrm{GL}_n(\mathbb{K})$. Pour k entier entre* 1 *et n, on note Δ_k le déterminant de la sous-matrice $g_k = g_{[1,k],[1,k]}$ formée des k premières lignes et des k premières colonnes de g.*
   (a) *Montrer que les conditions suivantes sont équivalentes :*
      (i) *il existe $\ell \in L$ et $b \in B$ telles que $g = \ell b$ (on dit que la méthode LU fonctionne sur g) ;*
      (ii) *les déterminants $\Delta_1, \ldots, \Delta_n$ ne sont pas nuls.*
   (b) *Montrer que si ces conditions sont satisfaites, le couple (ℓ, b) est unique.*
   (c) *Montrer qu'il existe une unique matrice de permutation π telle que $\pi^{-1} g$ satisfasse aux conditions (i) et (ii).*
   (d) *Évaluer la complexité du calcul de π, ℓ et b en fonction de n.*
3. *On prend $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Montrer que l'ensemble des matrices pour lesquelles la méthode LU échoue est la réunion d'un nombre fini de sous-variétés[8] et que c'est un fermé. [C'est l'adhérence de toutes les cellules de Bruhat, sauf une.]*
4. *On prend pour $\mathbb{K}$ un corps fini de cardinal q. Montrer que le nombre de matrices qui admettent une décomposition ℓb avec $\ell \in L$, $b \in B$, est $(q-1)^n q^{n(n-1)}$. Comparer ce nombre au cardinal de $\mathrm{GL}_n(\mathbb{F}_q)$; en particulier, donner un équivalent de la différence, lorsque q ou n est grand.*

B.5. Exercice. *Soit E un espace vectoriel de dimension finie et F un sous-espace. On note G_F le stabilisateur de F dans $\mathrm{GL}(E)$:*

$$G_F := \{g \in \mathrm{GL}(E),\ g(F) \subset F\}.$$

1. *Montrer que tout élément g de G_F définit par restriction un élément g_F de $\mathrm{GL}(F)$ et par passage au quotient un élément g^F de $\mathrm{GL}(E/F)$.*

[8] Ici, une sous-variété doit être comprise comme l'ensemble des zéros d'une famille finie de polynômes en n variables dans un espace vectoriel de dimension n.

2. *Montrer que l'application* π *de* G_F *vers le produit direct* $\mathrm{GL}(F)\times\mathrm{GL}(E/F)$, *qui à* g *associe* (g_F, g^F) *est un morphisme surjectif de groupes.*
3. *Montrer que tout supplémentaire de* F *définit une section de* π.
4. *Soit* h *dans* $\operatorname{Ker}\pi$. *Montrer que* $h - \mathrm{Id}_E$ *définit par passage au quotient un élément* φ_h *de* $\hom(E/F, F)$ *et que* $\varphi : \operatorname{Ker}\pi \to \hom(E/F, F)$, $\varphi(h) = \varphi_h$ *est un morphisme de groupes.*
5. *En déduire que l'on a un isomorphisme entre* G_F *et un produit semi-direct* $\hom(E/F, F) \rtimes (\mathrm{GL}(F) \times \mathrm{GL}(E/F))$ *que l'on précisera.*

B.6. Exercice. *Soit* $(e_i)_{1\leqslant i\leqslant n}$ *la base canonique de* $\mathbb{R}^n$ *muni de sa forme euclidienne habituelle. On considère la projection orthogonale* π *sur le sous-espace* $F = \langle e_i\rangle_{1\leqslant i\leqslant d}$.

1. *Montrer que* $g \mapsto g\pi g^{-1}$ *définit une application continue et surjective de* $\mathrm{O}(n)$ *sur l'ensemble* $\mathscr{P}_d$ *des projecteurs orthogonaux de rang* d *de l'espace* $\mathbb{R}^n$.
2. *En déduire un homéomorphisme entre* $\mathrm{Gr}(d, n)$ *et* $\mathscr{P}_d$.

Chapitre V

Groupes conservant une forme bilinéaire

Nous allons maintenant présenter les groupes classiques fixant une forme bilinéaire donnée. Encore une fois, le déroulement du chapitre passe par les actions de groupes et invariants d'actions. Tout d'abord, la classification des formes bilinéaires symétriques se ramène à l'étude des orbites du groupe linéaire $\mathrm{GL}_n(\mathbb{K})$ pour l'action par congruence. On dégagera des invariants partiels (rang, discriminant), puis des invariants totaux, permettant de caractériser les orbites. Mais un invariant, qu'il soit partiel ou total, dépend dramatiquement du corps choisi. Ici, on se limitera aux corps $\mathbb{R}$, $\mathbb{C}$, $\mathbb{F}_q$; pour aller plus loin, il est conseillé de feuilleter le *Cours d'arithmétique* [34] de Jean-Pierre Serre ou le futur classique sur les formes quadratiques : le livre de Clément de Seguins-Pazzis [33].

Une fois la classification faite, à l'aide de ces invariants totaux, les groupes classiques étudiés arriveront naturellement comme stabilisateurs d'éléments choisis dans une orbite (on rappelle que le stabilisateur ne dépend pas, à isomorphisme près de l'élément choisi dans une orbite). On trouvera en annexe A, la plupart des définitions, notations, et résultats classiques utilisés dans ce chapitre.

1. Classification des formes quadratiques en termes d'actions par congruence

Pour des raisons qui paraîtront évidentes par la suite, on travaille sur un corps $\mathbb{K}$ de caractéristique différente de 2. On travaille dans $\mathbb{K}^n$, où n est

un entier fixé. Le bon objet d'étude tout au long de cette section est, pour le goût personnel de la moitié des auteurs, la forme quadratique, c'est-à-dire un polynôme homogène de degré 2 (en les coordonnées, dans une base quelconque). Appelons q une telle forme, voir aussi la définition B.1.2. Il existe une pléthore de formes bilinéaires b telles que $b(x,x) = q(x)$ pour tout vecteur x. Mais il en existe une unique qui soit symétrique, il s'agit de

$$b(x,y) = \frac{q(x+y) - q(x) - q(y)}{2}.$$

Le fait de substituer l'étude d'une forme bilinéaire à cette d'une forme quadratique est astucieuse, en fait, selon la stratégie des Horaces et des Curiaces, on préfère attaquer deux fois le degré 1 (« bi-linéaire ») plutôt qu'une fois le degré 2 (quadratique). Enfin, le choix d'une forme bilinéaire *symétrique* provient de l'unicité ci-dessus.

Tous les outils matriciels pour aborder les formes quadratiques sont rappelés en annexe A.

1.1. Définition (Congruence). Soit $\mathbb{K}$ un corps et n un entier naturel. L'action de *congruence* est l'action de $\mathrm{GL}_n(\mathbb{K})$ sur $\mathscr{S}_n(\mathbb{K})$, l'espace des matrices symétriques $n \times n$, définie par :

$$\forall g \in \mathrm{GL}_n(\mathbb{K}),\ \forall A \in \mathscr{S}_n(\mathbb{C}), \quad g \cdot A := gA\,{}^t g.$$

On dit que deux matrices symétriques A et A' de $\mathscr{S}_n(\mathbb{K})$ sont *congruentes* si elles appartiennent à la même orbite pour l'action de congruence.

Il faut retenir que deux matrices sont congruentes si et seulement si elles sont les matrices de la même forme quadratique dans des bases différentes. Dans la suite, on notera $\mathrm{Orb}(A)$ l'orbite de congruence de A et le but est maintenant de trouver des invariants raisonnables, si possible totaux, permettant de déterminer si deux matrices sont ou non dans la même orbite, c'est-à-dire si elles sont ou non congruentes.

Les invariants rang et discriminant, sont définis en A.1.7 et A.1.9. Pour faire bref, le rang de la forme est égal au rang de la matrice associée dans une base quelconque et le discriminant est égal au déterminant de la matrice modulo les carrés (non nuls), et ils sont invariants pour l'action de congruence puisqu'ils ne dépendent pas de la base choisie. On les note $\mathrm{rg}(A)$ et $\delta(A)$ ou, comme ils ne dépendent que de b et pas de A : $\mathrm{rg}(b)$ et $\delta(b)$. Finesse : si la forme bilinéaire b associée à A est dégénérée, on remplace $\delta(A)$ par $\delta(\overline{A})$, le discriminant de la forme bilinéaire sur $\mathbb{K}^n/\operatorname{Ker} b$ (voir le § A.1).

Dans le cas réel, on peut définir un nouvel invariant :

$$p(b) := \max\{\dim F \,:\, F \subset E = \mathbb{K}^n,\ \forall x \in F, x \neq 0 \Rightarrow\ q(x) > 0\}.$$

L'entier $p(b)$ ne dépend que de la forme symétrique bilinéaire b (ou si l'on préfère, de la forme quadratique q) et non de A, donc il ne dépend pas de la base choisie. Il s'agit bien d'un invariant de l'action. On pourra le noter $p(A)$ si A est la matrice associée à b dans une base.

Le théorème suivant montre que ces invariants peuvent être des invariants totaux dans le cas des corps classiques $\mathbb{C}$, $\mathbb{R}$ et $\mathbb{F}_q$.

1.2. Théorème

(i) *Si* $\mathbb{K} = \mathbb{C}$, *deux matrices* $A, A' \in \mathscr{S}_n(\mathbb{C})$ *sont dans la même orbite si et seulement si elles ont le même rang :*

$$\mathrm{Orb}(A) = \mathrm{Orb}(A') \Longleftrightarrow \mathrm{rg}\, A = \mathrm{rg}\, A'.$$

(ii) *Si* $\mathbb{K} = \mathbb{R}$, *deux matrices* $A, A' \in \mathscr{S}_n(\mathbb{R})$ *sont dans la même orbite si et seulement si elles ont le même rang et même invariant* p *:*

$$\mathrm{Orb}(A) = \mathrm{Orb}(A') \Longleftrightarrow \mathrm{rg}\, A = \mathrm{rg}\, A' \text{ et } p(A) = p(A').$$

(iii) *Si* $\mathbb{K} = \mathbb{F}_q$ *est un corps fini de cardinal impair, deux matrices* inversibles $A, A' \in \mathscr{S}_n(\mathbb{F}_q)$ *sont dans la même orbite si et seulement si elles ont le même discriminant :*

$$\mathrm{Orb}(A) = \mathrm{Orb}(A') \Longleftrightarrow \delta(A) = \delta(A').$$

Démonstration. On a déjà vu que les quantités considérées (rang, discriminant, « invariant p ») sont invariantes, il suffit donc de montrer dans chaque cas que toutes matrices A ayant ces invariants sont congruentes à une matrice particulière, qui fait office de *forme normale* dans l'orbite. Pour cela, on utilise un procédé d'orthogonalisation par récurrence descendante sur la dimension.

(i) $\boxed{\mathbb{K} = \mathbb{C}}$. Soit $A \in \mathscr{S}_n(\mathbb{C})$, avec $\mathrm{rg}\, A = r \neq 0$ (si $r = 0$, l'assertion est triviale). Alors, il existe un vecteur u tel que $q(u) \neq 0$ (sinon b serait nulle). Soit $\lambda \in \mathbb{C}^*$ tel que $\lambda^2 = q(u)$ et $v := u/\lambda$. Alors : $q(v) = 1$. La forme b étant non dégénérée sur $\mathbb{C}v$, on a $E = \mathbb{C}v \oplus (\mathbb{C}v)^\perp$ par le lemme A.1.13.

Donc A est congruente à une matrice de la forme

$$\begin{pmatrix} 1 & 0 \\ 0 & A' \end{pmatrix},$$

où $\mathrm{rg}\, A' = r - 1$. Par itération, on montre que A est congruente à

$$\begin{pmatrix} \mathrm{I}_k & 0 \\ 0 & 0 \end{pmatrix},$$

d'où, par invariance du rang, on a nécessairement : $k = r$.

(ii) $\boxed{\mathbb{K} = \mathbb{R}}$ Soit $A \in \mathscr{S}_n(\mathbb{R})$ telle que $\operatorname{rg} A = r$, $p(A) = p$ et soit $q = n - p - r$.

La réduction est analogue au cas complexe sauf que...
– si $q(u) > 0$, on pose $\lambda = q(u)^{-1/2}$ et $v = \lambda u$, d'où : $q(v) = 1$;
– si $q(u) < 0$, on pose $\lambda = |q(u)|^{-1/2}$ et $v = \lambda u$ d'où : $q(v) = -1$.
Alors, en itérant le procédé, puis en permutant les vecteurs de la base pour regrouper les coefficients par signe, on voit que A est congruente à :

$$A' := \begin{pmatrix} \mathrm{I}_{p'} & & \\ & -\mathrm{I}_{q'} & \\ & & 0 \end{pmatrix}.$$

Or, on a : $p(A') = p'$. Cela résulte en effet de la formule de Grassmann : si F' est un sous-espace de dimension $d \geqslant p' + 1$, l'intersection de F' avec le sous-espace G' engendré par les $n - p'$ derniers vecteurs de la base dans laquelle b a pour matrice A' est de dimension

$$\dim(F' \cap G') = \dim F' + \dim G' - \dim(F' + G') \geqslant p' + 1 + n - p' - n = 1,$$

si bien que $F' \cap G'$ contient un vecteur non nul v tel que $q(v) \leqslant 0$. Ainsi, la restriction de la forme quadratique à F' n'est pas définie positive et $p(A') \leqslant p'$. L'inégalité $p(A') \geqslant p'$ est claire grâce à l'espace engendré par les p' premiers vecteurs de base.

Alors, par invariance de « l'invariant p », on a : $p' = p$, puis, comme le rang de A est $p + q = p' + q'$, la conclusion : $q = q'$.

(iii) $\boxed{\mathbb{K} = \mathbb{F}_q}$ On suppose que A est inversible, *i.e.* la forme bilinéaire associée est non dégénérée. Soit ζ un non-carré de $\mathbb{F}_q^*$.

En utilisant le même procédé d'orthogonalisation que précédemment, on voit que A est congruente à

$$\begin{pmatrix} \epsilon_1 & & \\ & \ddots & \\ & & \epsilon_n \end{pmatrix},$$

avec $\epsilon_i = 1$ ou ζ. On peut encore faire mieux. Pour cela, on a besoin de l'exercice très classique suivant.

Exercice. *Montrer que l'équation $ax^2 + by^2 = 1$, où $a, b \neq 0$, admet au moins une solution dans $\mathbb{F}_q \times \mathbb{F}_q \setminus \{(0,0)\}$.*

[Montrer qu'il existe $(q+1)/2$ éléments de la forme ax^2 et autant de la forme $1 - by^2$. Ainsi, puisque $(q+1)/2 + (q+1)/2 > q$, il existe un élément de la forme $ax^2 = 1 - by^2$.]

D'après l'exercice, dans un espace de dimension $n \geqslant 2$, toute forme quadratique admet un vecteur v tel que $q(v) = 1$. On déduit alors que A est congruente à

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \epsilon_n \end{pmatrix}.$$

(En effet, soit on reprend le procédé d'orthogonalisation en imposant grâce à l'exercice qu'à chaque étape, sauf peut-être la dernière (où l'on n'a plus qu'un espace de dimension 1), au vecteur v la relation $q(v) = 1$; soit on remarque que les formes $\zeta x^2 + \zeta y^2$ et $x^2 + y^2$ sont congruentes (prendre une base (v, v') de $\mathbb{F}_q^2$ où $q(v) = 1$ et v' orthogonal à v), si bien que l'on peut avoir au plus un coefficient ϵ_i égal à ζ). □

1.3. Remarque. Le cas $\mathbb{K} = \mathbb{R}$ est une formulation du théorème d'inertie de Sylvester, qui est passé en revue en annexe, voir théorème B.3.4 et ses différentes versions.

Notons aussi que sur tout corps de caractéristique différente de 2, toute matrice est congruente à une matrice diagonale, voir aussi la proposition B.3.1 en annexe, pour ce résultat.

1.4. Remarque. Pour les corps finis, la classification ne porte que sur les orbites de matrices symétriques non dégénérées. L'énoncé complet serait le suivant : sur l'espace $\mathbb{F}_q^n$, q impair, une forme bilinéaire symétrique b est caractérisée par son rang et par le discriminant réduit $\bar{\delta}(b)$, qui est le discriminant de la forme quadratique obtenue en quotientant $\mathbb{F}_q^n$ par le noyau de b (paragraphe A.1).

En effet, soit A une matrice symétrique de rang r, b_A la forme bilinéaire associée sur $\mathbb{F}_q^n$, $\bar{\delta}(A) = \bar{\delta}(b_A) \in \{1, \zeta\}$. Alors, A est congruente à la matrice diagonale dont les coefficients diagonaux sont $(0, \ldots, 0, 1, \ldots, 1, \varepsilon)$ où il y a $n-r$ coefficients nuls et $\varepsilon \in \{1, \zeta\}$ est congruente à $\bar{\delta}(A)$ modulo les carrés. Pour le voir, compléter une base $\mathbf{e}$ du noyau par une base $\mathbf{e}'$ d'un supplémentaire et réduire b_A sur ce supplémentaire en utilisant la proposition.

Il est alors naturel de s'intéresser au groupe d'isotropie de chaque orbite. Ces groupes sont présentés matriciellement comme stabilisateurs d'une matrice, et donc, à isomorphisme près, comme stabilisateur de tout élément de l'orbite correspondante. Mais, une fois une base fixée, ils correspondent au groupe des isomorphismes qui conserve la forme bilinéaire attachée à la matrice.

$\boxed{\mathbb{K} = \mathbb{C}}$ Toute orbite de congruence rencontre donc une matrice de la forme

$$\mathrm{I}_{n,n,r} = \begin{pmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Il est donc tout aussi naturel de s'intéresser aux orbites :

$$\mathrm{Orb}(\mathrm{I}_n) = \{P\,{}^tP \in \mathscr{M}_n(\mathbb{C}) \,:\, P \in \mathrm{GL}_n(\mathbb{C})\}$$
$$\mathrm{Orb}(\mathrm{I}_{n,n,r}) = \{P\mathrm{I}_{n,n,r}\,{}^tP \in \mathscr{M}_n(\mathbb{C}) \,:\, P \in \mathrm{GL}_n(\mathbb{C})\}$$

et à leurs groupes d'isotropie respectifs.

On définit le groupe d'isotropie de I_n :

$$\begin{aligned}\mathrm{O}(n,\mathbb{C}) = \mathrm{Stab}(\mathrm{I}_n) &= \{P \in \mathrm{GL}_n(\mathbb{C}) \,:\, P\,{}^tP = \mathrm{I}_n\} \\ &= \{P \in \mathrm{GL}_n(\mathbb{C}) \,:\, {}^tPP = \mathrm{I}_n\}.\end{aligned}$$

(La dernière égalité vient du fait que $\mathrm{GL}_n(\mathbb{C})$ est stable par la transposée.)

On définit le groupe d'isotropie de $\mathrm{I}_{n,n,r}$:

$$\mathrm{Stab}(\mathrm{I}_{n,n,r}) = \mathrm{O}(n,r,\mathbb{C}) := \{P \in \mathrm{GL}_n(\mathbb{C}) \,:\, P\mathrm{I}_{n,n,r}\,{}^tP = \mathrm{I}_{n,n,r}\}.$$

1.5. Exercice (Dévissage de $\mathrm{O}(n,r,\mathbb{C})$). *Montrer que*

$$\mathrm{O}(n,r,\mathbb{C}) \simeq \big(\mathrm{O}(r,\mathbb{C}) \times \mathrm{GL}_{n-r}(\mathbb{C})\big) \ltimes_\varphi \mathscr{M}_{r,n}(\mathbb{C}),$$

où, avec des notations évidentes, $\varphi : (A,D)\cdot C \longmapsto ACD^{-1}$.

Un nouveau groupe « classique » important est donc $\mathrm{O}(n,\mathbb{C})$ puisqu'il permet de comprendre les autres groupes d'isotropie.

$\boxed{\mathbb{K} = \mathbb{R}}$ Toute orbite de congruence rencontre donc une matrice de la forme

$$\mathrm{I}_{(p,q)} = \begin{pmatrix} \mathrm{I}_p & 0 & 0 \\ 0 & -\mathrm{I}_q & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

On s'intéresse donc à leurs orbites, en s'arrêtant un peu sur le cas particulier de l'orbite de l'identité. On pose :

$$\mathscr{S}_n^{++}(\mathbb{R}) := \mathrm{Orb}(\mathrm{I}_n) = \{P\,{}^tP \in \mathscr{M}_n(\mathbb{R}) \,:\, P \in \mathrm{GL}_n(\mathbb{R})\}.$$

Aparté. Vu comme application bilinéaire symétrique, c'est l'ensemble des *produits scalaires.* Matriciellement, il s'agit des matrices symétriques *définies positives.* La raison de cette définition (et de cette notation) réside dans le résultat suivant.

1.6. Proposition. *L'orbite de l'identité pour l'action de congruence réelle,*

$$\{P\,{}^tP \in \mathscr{M}_n(\mathbb{R}) \,:\, P \in \mathrm{GL}_n(\mathbb{R})\},$$

peut aussi s'écrire

$$\{S \in \mathrm{GL}_n(\mathbb{R}) \,:\, {}^tS = S;\ \forall x \in \mathbb{R}^n \setminus \{0\},\ {}^txSx > 0\}.$$

Démonstration. L'inclusion du premier dans le second est claire puisque ${}^txP\,{}^tPx = {}^t({}^tPx)({}^tPx)$, qui est une somme de carrés, donc positif et nul si et seulement si ${}^tPx$ est le vecteur nul, c'est-à-dire si x est le vecteur nul.

Pour la réciproque, on sait que chaque orbite de congruence rencontre une matrice de la forme $\mathrm{I}_{(p,q)}$. Donc, si S vérifie ${}^txSx > 0$ pour tout x non nul, tous les termes diagonaux doivent être strictement positifs puisqu'ils sont égaux à ${}^te_iSe_i$, où (e_i) est la base canonique de $\mathbb{R}^n$. Ainsi, $p = n$ et $q = 0$, et donc S est congruente à $\mathrm{I}_{(p,q)} = \mathrm{I}_n$. □

1.7. Définition. On appelle *groupe orthogonal réel d'ordre n* le groupe d'isotropie de la matrice symétrique réelle I_n :

$$\mathrm{Stab}(\mathrm{I}_n) = \mathrm{O}(n,\mathbb{R}) = \{P \in \mathrm{GL}_n(\mathbb{R}) \,:\, P\,{}^tP = \mathrm{I}_n\}.$$

Ce groupe sera noté tout simplement $\mathrm{O}(n)$.

On s'intéresse ensuite à l'orbite des matrices symétriques de signature (p,q), voir aussi en annexe la définition B.3.5 de la signature :

$$\mathrm{Orb}(\mathrm{I}_{(p,q)}) = \{P\mathrm{I}_{(p,q)}\,{}^tP \in \mathscr{M}_n(\mathbb{R}) \,:\, P \in \mathrm{GL}_n(\mathbb{R})\}.$$

On définit le groupe d'isotropie de $\mathrm{I}_{(p,q)}$: $\mathrm{Stab}(\mathrm{I}_{(p,q)}) = \{P \in \mathrm{GL}_n(\mathbb{R}) : P\mathrm{I}_{(p,q)}\,{}^tP = \mathrm{I}_{(p,q)}\}$. Lorsque $\mathrm{I}_{(p,q)}$ est de rang n, c'est-à-dire si $p+q=n$, on notera ce groupe $\mathrm{O}(p,q)$. On en trouve une étude topologique (détermination de ses composantes connexes) en application de la décomposition polaire au paragraphe VI-A.

De la même manière que dans le cas complexe, tous les groupes d'isotropie se ramènent au groupe $\mathrm{O}(n,\mathbb{C})$, on peut, dans le cas réel, ramener l'étude des groupes d'isotropie à celle des groupes $\mathrm{O}(n)$ et $\mathrm{O}(p,q)$. C'est devenu maintenant un réflexe de s'intéresser à l'ordre de dégénérescence. Le résultat suivant constitue un exercice qui se fait hop-hop en deux inclusions.

1.8. Exercice. *Montrer que* $\overline{\mathrm{Orb}(\mathrm{I}_{(p,q)})} = \bigsqcup_{\substack{0\leqslant h\leqslant p\\ 0\leqslant h'\leqslant q}} \mathrm{Orb}(\mathrm{I}_{(p-h,q-h')})$.

1.9. Exercice. *À l'aide du théorème précédent, décrire l'intersection d'une orbite pour la congruence complexe dans $\mathscr{M}_n(\mathbb{C})$ avec $\mathscr{M}_n(\mathbb{R})$ comme une union d'orbites pour la congruence réelle. Montrer par exemple :*

$$\mathrm{GL}_n(\mathbb{R}) \cap \mathrm{GL}_n(\mathbb{C}) \cdot \mathrm{I}_n = \bigcup_{p+q=n} \mathrm{GL}_n(\mathbb{R}) \cdot \mathrm{I}_{(p,q)}.$$

Pour l'inclusion, utiliser le théorème et un argument sur le rang. Pour l'inclusion inverse montrer que $\mathrm{I}_{(p,q)} \in \mathrm{GL}_n(\mathbb{R}) \cap \mathrm{GL}_n(\mathbb{C}) \cdot \mathrm{I}_n$, puis conclure.

1.10. Exercice. *Généraliser ce résultat pour les autres orbites de congruence complexe.*

$\boxed{\mathbb{K} = \mathbb{F}_q}$ Il y a deux orbites de matrices non dégénérées. Pour une description des autres orbites, voir l'exercice 1.4.

$$\mathrm{Orb}(\mathrm{I}_n) = \{P\,{}^tP \in \mathscr{M}_n(\mathbb{F}_q) \,:\, P \in \mathrm{GL}_n(\mathbb{F}_q)\},$$
$$\mathrm{Orb}(\mathrm{I}_\zeta) = \{P\mathrm{I}_\zeta\,{}^tP \in \mathscr{M}_n(\mathbb{F}_q) \,:\, P \in \mathrm{GL}_n(\mathbb{F}_q)\},$$
$$\text{où } \mathrm{I}_\zeta = \begin{pmatrix} \mathrm{I}_{n-1} & 0 \\ 0 & \zeta \end{pmatrix} \quad \text{et} \quad \zeta \text{ un non-carré.}$$

1.11. Définition. Il y a deux types de *groupes orthogonaux finis* d'ordre n : le groupe d'isotropie de I_n :

$$\mathrm{O}(n, \mathbb{F}_q) = \mathrm{Stab}(\mathrm{I}_n) = \{P \in \mathrm{GL}_n(\mathbb{F}_q) \,:\, P\,{}^tP = \mathrm{I}_n\}.$$

et le groupe d'isotropie de I_ζ :

$$\mathrm{O}^\zeta(n, \mathbb{F}_q) = \mathrm{Stab}(\mathrm{I}_\zeta) = \{P \in \mathrm{GL}_n(\mathbb{F}_q) \,:\, P\mathrm{I}_\zeta\,{}^tP = \mathrm{I}_\zeta\}.$$

2. Le cas hermitien

Lorsque l'on travaille en complexe, le groupe $\mathrm{O}(n, \mathbb{C})$ peut bien entendu intervenir, mais on lui préférera le groupe $\mathrm{U}(n)$ qui a le bon goût d'être compact. Nous passons rapidement en revue les résultats, très similaires à ceux qui précèdent, qui permettent de l'introduire.

2.1. Définition. Soit $P \in \mathscr{M}_n(\mathbb{C})$ une matrice complexe. Sa *matrice adjointe* est la transconjuguée[1] $P^* := {}^t\overline{P}$. Une matrice est dite *hermitienne* si elle est égale à sa transconjuguée.

[1] Élégante contraction de *transposée* et *conjuguée*.

Tout comme les matrices symétriques codent des formes bilinéaires symétriques, les matrices hermitiennes codent des formes sesquilinéaires.

Le cas hermitien est analogue au cas euclidien (réel) et laissé en exercice, on introduit l'action par congruence hermitienne :

$$\begin{array}{ccc} \mathrm{GL}_n(\mathbb{C}) \times \mathscr{M}_n(\mathbb{C}) & \longrightarrow & \mathscr{M}_n(\mathbb{C}) \\ (P, A) & \longmapsto & PAP^*. \end{array}$$

La signature ainsi que le rang classent les orbites. L'ensemble des produits scalaires hermitiens[2] est :

$$\mathscr{H}_n^{++}(\mathbb{C}) = \mathrm{Orb}(\mathrm{I}_n) = \{PP^* \in \mathscr{M}_n(\mathbb{C}) \,:\, P \in \mathrm{GL}_n(\mathbb{C})\}.$$

Les autres orbites sont classées par les paires d'entiers (p, q), $p + q \leqslant n$:

$$\mathrm{Orb}(\mathrm{I}_{(p,q)}) = \{P\mathrm{I}_{(p,q)}P^* \in \mathscr{M}_n(\mathbb{C}) \,:\, P \in \mathrm{GL}_n(\mathbb{C})\}.$$

Comme dans le cas symétrique réel, on met à part l'orbite de l'identité, qui est aussi, voir par analogie la proposition 1.6, l'ensemble des matrices *hermitiennes définies positives*, ou l'ensemble des *produits scalaires hermitiens*.

2.2. Définition. Soit n un entier. Le *groupe unitaire* d'ordre n est le groupe d'isotropie de I_n :

$$\mathrm{U}(n) = \mathrm{Stab}(\mathrm{I}_n) = \{P \in \mathrm{GL}_n(\mathbb{C}) \,:\, PP^* = \mathrm{I}_n\}.$$

Soit p et q deux entiers tels que $p + q = n$. Le groupe d'isotropie de la matrice $\mathrm{I}_{(p,q)}$ est noté :

$$\mathrm{U}(p, q) = \mathrm{Stab}(\mathrm{I}_{(p,q)}) = \{P \in \mathrm{GL}_n(\mathbb{C}) \,:\, P\mathrm{I}_{(p,q)}P^* = \mathrm{I}_{(p,q)}\}.$$

3. Les groupes orthogonaux et leurs actions

Soit q une forme quadratique sur $\mathbb{K}$ et b sa forme bilinéaire symétrique associée de matrice A. On associe à q son groupe d'isotropie.

3.1. Définition. On définit le groupe d'isotropie de la forme quadratique q ou de sa forme bilinéaire associée b par

$$\mathrm{O}(q) = \mathrm{O}(b) = \mathrm{Stab}(A) = \{P \in \mathrm{GL}_n(\mathbb{K}) \,:\, PA\,{}^tP = A\}.$$

C'est aussi le groupe des matrices $P \in \mathrm{GL}_n(\mathbb{K})$ telles que pour tous $x, y \in \mathbb{K}^n$, $b(Px, Py) = b(x, y)$. Un élément de $\mathrm{O}(q)$ est appelé *isométrie* de q.

[2]Remarque : ce sont les *matrices hermitiennes définies positives*. On verra que cet ensemble est homéomorphe à l'espace vectoriel réel des matrices hermitiennes (de dimension n^2) via l'exponentielle de matrices, p. 208.

On retrouve donc parmi ces groupes les groupes classiques $\mathrm{O}(n,\mathbb{C})$, $\mathrm{O}(n)$, $\mathrm{O}(p,q)$, $\mathrm{O}^{\varsigma}(n,\mathbb{F}_q),\ldots$ vus précédemment et dont l'intérêt vient de la classification des orbites de congruence. Plus généralement, on notera $\mathrm{O}(n,\mathbb{K})$ le stabilisateur de l'identité dans $\mathrm{GL}_n(\mathbb{K})$ pour la congruence.

3.2. Remarque. Si deux formes quadratiques q et q' sont dans la même orbite pour l'action de congruence de $\mathrm{GL}_n(\mathbb{K})$, alors $\mathrm{O}(q)$ et $\mathrm{O}(q')$ sont isomorphes (et même conjugués). On peut se demander si la réciproque est vraie. On va voir qu'il n'en est rien.

On peut faire la remarque naïve que si $PA\,{}^tP = A$, alors $P(-A)\,{}^tP = -A$. On en déduit facilement que $\mathrm{O}(p,q) = \mathrm{O}(q,p)$. Ce qui compromet une réciproque : on peut très bien avoir des stabilisateurs isomorphes (et même carrément égaux) sans que les matrices soient dans la même orbite.

L'expérience que nous avons acquise sur les groupes montre que leur étude passe par leurs actions naturelles. Sur quoi $\mathrm{O}(q)$ agit-il naturellement ? Nous avons vu que $\mathrm{O}(n)$ agit sur la sphère $\mathbf{S}^{n-1}$. De la même manière, la définition que $\mathrm{O}(q)$ assure qu'il agit sur les nappes, appelées *quadriques*, $N_k = \{x \in \mathbb{K}^n \,:\, q(x) = k\}$, où $k \in \mathbb{K}$.

En effet, on voit que pour tout P de $\mathrm{O}(b)$ et tout $x \in N_k$, on a : $q(Px) = b(Px,Px) = b(x,x) = q(x) = k$, ce qui signifie que $Px \in N_k$. Les nappes constituent aussi un objet d'étude important en soi.

3.3. Question. Cette action est-elle transitive ?

Regardons sur un exemple. Soit E l'espace vectoriel euclidien de dimension n, et N_1 la sphère de rayon 1. On sait, d'après le théorème de la base orthonormée incomplète, que l'action est transitive. Effectivement, tout vecteur de norme 1 peut être complété en une base orthonormée et l'action (simplement) transitive de $\mathrm{O}(n)$ sur l'ensemble des bases orthonormées implique alors la transitivité de l'action de $\mathrm{O}(n)$ sur la sphère. La transitivité sur N_k en découle.

On peut aussi se poser la question de transitivité suivante (pour q quelconque) : l'action de $\mathrm{O}(q)$ est elle transitive sur une grassmannienne ? Dit autrement, pour tout $F, F' \subset E$ de même dimension, existe-t-il $P \in \mathrm{O}(q)$ tel que $P(F) = F'$? Il est clair que la réponse est non en général : prendre pour F une droite engendrée par un vecteur v isotrope, *i.e.* tel que $q(v) = 0$, et pour F' une droite engendrée par un vecteur non isotrope. Par exemple, pour la forme réelle, de signature $(1,1)$, donnée par $q(x,y) = x^2 - y^2$, la droite engendrée par $(1,1)$ ne peut être envoyée par $\mathrm{O}(q)$ sur la droite engendrée par $(1,0)$.

Les théorèmes de base incomplète, si utiles jusqu'à présent, sont dans le cadre d'un espace muni d'une forme quadratique, remplacés par le puissant théorème de Witt, et dont une preuve peut être lue dans le livre [29], mais que l'on peut trouver dans l'exercice D.14.

3.4. Théorème (Théorème de Witt). *Soit un espace vectoriel E muni d'une forme bilinéaire non dégénérée b sur $\mathbb{K}$ de caractéristique différente de 2, et $F, F' \subset E$ deux sous-espaces. Alors, toute isométrie $\sigma : F \to F'$ peut-être prolongée en une isométrie $\tilde{\sigma} : E \to E$.*

Cela répond au problème de transitivité sur les nappes.

3.5. Corollaire. *Pour tout scalaire k non nul, le groupe orthogonal $\mathrm{O}(q)$ agit transitivement sur la nappe $N_k = \{v,\ q(v) = k\}$. La nappe N_0 possède (au plus) deux orbites : $\{0\}$ et $N_0 \backslash \{0\}$.*

Démonstration. Ceci est quasiment immédiat si k est non nul. Soit v et v' tels que $q(v) = k = q(v')$. Comme k est non nul, $\mathbb{K}v$ et $\mathbb{K}v'$ sont des droites et de plus, les hypothèses impliquent que l'application $\alpha v \mapsto \alpha v'$ est une isométrie de $\mathbb{K}v$ et $\mathbb{K}v'$. On prolonge cette isométrie en isométrie de tout l'espace par le théorème de Witt, ce qui assure un élément g de $O(q)$ tel que $g(v) = v'$.

Si $k = 0$, la méthode est la même dès que l'on choisit v et v' non nuls. □

Un peu plus général, le théorème de Witt peut être vu comme un théorème de multi-transitivité (transitivité sur des m-uplets) sous des conditions métriques.

3.6. Corollaire. *Soit b une forme bilinéaire symétrique sur $\mathbb{K}^n$ et soit $(x_i)_{1 \leqslant i \leqslant m}$ et $(y_j)_{1 \leqslant j \leqslant m}$ deux familles libres de $\mathbb{K}^n$ telles que l'on ait : $b(x_i, x_j) = b(y_i, y_j)$ pour tout (i, j).*

Alors, il existe une isométrie $P \in \mathrm{O}(b)$ telle que $Px_i = y_i$ pour tout (i, j).

Voici un exemple important à avoir en tête.

3.7. Exemple (Angles du plan). Dans le plan réel euclidien, le groupe des isométries $\mathrm{O}(2)$ et même le groupe des rotations $\mathrm{SO}(2)$ agit sur les couples de vecteurs du cercle $\mathbf{S}^1$ (cas $m = 2$). La condition $b(x_1, x_2) = b(y_1, y_2)$ traduit l'égalité des angles : $\cos(x_1, x_2) = \cos(y_1, y_2)$. Le lecteur pourra en déduire que le quotient $(\mathbf{S}^1 \times \mathbf{S}^1)/\mathrm{SO}(2)$ s'identifie à l'ensemble des angles orientés et le quotient $(\mathbf{S}^1 \times \mathbf{S}^1)/\mathrm{O}(2)$ aux angles non orientés. Il s'agit là d'un exemple à retenir où les groupes permettent de reconstruire la géométrie classique.

4. Formes bilinéaires antisymétriques

Il n'y a qu'à considérer l'importance primordiale du déterminant pour comprendre la profondeur de la notion des k-formes alternées (ou antisymétriques). Les formes bilinéaires antisymétriques gardent une importance bien à elles que nous n'aurons pas le loisir d'observer dans cet ouvrage. L'omniprésence de structures dites symplectiques sur les variétés classiques a donné aux mathématiciens et aux physiciens des outils simples et puissants. Nous nous limiterons ici à l'introduction de quelques définitions, de quelques groupes importants dans ce domaine, et de quelques propriétés, tout en restant dans le cadre de l'algèbre linéaire. La plupart des résultats intéressants de ce livre concernant les formes antisymétriques sont présentés en exercice.

Soit un corps $\mathbb{K}$ (toujours de caractéristique différente de 2). Une forme bilinéaire b est dite antisymétrique si pour tous vecteurs $x, y \in \mathbb{K}^n$ on a : $b(x,y) = -b(y,x)$. On note

$$\mathscr{A}_n(\mathbb{K}) = \{A \in \mathscr{M}_n(\mathbb{K}) \,:\, {}^tA = -A\}$$

l'espace vectoriel des matrices antisymétriques. Le choix d'une base induit un isomorphisme entre l'espace des formes antisymétriques et l'espace $\mathscr{A}_n$.

4.1. Proposition. *Le groupe* $\mathrm{GL}_n(\mathbb{K})$ *agit par congruence sur* $\mathscr{A}_n(\mathbb{K})$, *continûment si cela a un sens, et les orbites sont classées par le rang, qui est pair.*

Démonstration. La première partie de la proposition est facile : ${}^t({}^tgAg) = -{}^tgAg$ si $g \in \mathrm{GL}_n(\mathbb{K})$ et $A \in \mathscr{A}_n(\mathbb{K})$ et la continuité est claire. Il s'agit donc de classer les orbites, ce qui résulte du lemme suivant. □

4.2. Lemme. *Soit E un espace vectoriel et b une forme bilinéaire antisymétrique sur E. Le rang de b est pair, et il existe une base de E dans laquelle la matrice de b est :*

$$J_{n,p} := \begin{pmatrix} 0 & \mathrm{I}_p & 0 \\ -\mathrm{I}_p & 0 & 0 \\ 0 & 0 & 0_{n-r} \end{pmatrix}, \quad \text{où} \quad \begin{cases} n = \dim E, \\ r = 2p = \operatorname{rg} b. \end{cases}$$

Démonstration. On procède par récurrence sur la dimension de E. Si E est réduit à $\{0\}$, l'assertion est vide. Si E est la droite engendrée par un vecteur v, alors b est nulle, car on a : $b(v,v) = -b(v,v)$; l'assertion du lemme en résulte dans ce cas.

Soit n un entier, supposons l'assertion du lemme vraie pour tout espace de dimension strictement inférieure à n. Soit b une forme non nulle (sinon, rien

à dire) sur un espace de dimension n. Il existe donc un couple de vecteurs (ϵ, ϵ') tel que $b(\epsilon, \epsilon') \neq 0$. Alors, (ϵ, ϵ') forme une famille libre, sinon l'on aurait $\epsilon' = k\epsilon$ pour $k \in \mathbb{K}$ convenable, d'où $b(\epsilon, \epsilon') = kb(\epsilon, \epsilon) = 0$.

On pose alors : $\epsilon_1 = b(\epsilon, \epsilon')^{-1}\epsilon$, $\epsilon_1' = \epsilon'$ et $F = \langle \epsilon, \epsilon' \rangle$: c'est un plan. La matrice de la restriction de b à F dans la base $(\epsilon_1, \epsilon_1')$ est :

$$\mathrm{mat}(b|_{F\times F}) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Elle est non dégénérée, ce qui permet d'écrire : $E = F \oplus F^\perp$. Ainsi, pour toute base $(e_3, \ldots, e_n)$ de $\mathbb{F}^\perp$, la matrice de b dans $(\epsilon_1, \epsilon_1', e_3, \ldots, e_n)$ est de la forme :

$$\begin{pmatrix} 0 & 1 & & \\ -1 & 0 & & \\ & & A' & \end{pmatrix},$$

où $A' \in \mathscr{A}_{n-2}(\mathbb{K})$ est la matrice de la restriction de b à $F^\perp$.

Par hypothèse de récurrence, on peut adjoindre à $(\epsilon_1, \epsilon_1')$ une base de $F^\perp$ de sorte que la matrice de b soit de la première forme souhaitée.

On obtient une base $(\epsilon_1, \epsilon_1', \epsilon_2, \epsilon_2', \ldots, \epsilon_p, \epsilon_p', \epsilon_{2p+1}, \ldots, \epsilon_n)$ que l'on réordonne en une base $(\epsilon_1, \epsilon_2, \ldots, \epsilon_p, \epsilon_1', \ldots, \epsilon_p', \epsilon_{2p+1}, \ldots, \epsilon_n)$, ce qui revient à permuter lignes et colonnes. On obtient alors une matrice de la forme voulue. □

4.3. Remarque. D'après la proposition, il ne peut exister de forme bilinéaire antisymétrique non dégénérée qu'en dimension paire. On peut le voir directement car si A est la matrice d'une telle forme, on a : $0 \neq \det A = \det {}^tA = \det(-A) = (-1)^n \det A$, ce qui force n à être pair.

4.4. Remarque. Une matrice antisymétrique est ainsi congruente à une matrice de type $J_{n,p}$. On montre aussi que, dans le cadre réel, elle peut être diagonalisée dans une base orthonormée et que ses valeurs propres sont imaginaires pures ; c'est une conséquence immédiate de l'exercice D.19.

Revenons à l'action de $\mathrm{GL}_n(\mathbb{K})$ sur $\mathscr{A}_n(\mathbb{K})$ par congruence. L'orbite de J_n est : $\mathrm{Orb}(J_n) = \{PJ_n\,{}^tP \in \mathscr{A}_n(\mathbb{K}) : P \in \mathrm{GL}_n(\mathbb{K})\}$.

4.5. Définition. Soit $n = 2p$ un entier pair. On appelle *groupe symplectique* de rang p le groupe d'isotropie de J_n :

$$\mathrm{Sp}_{2p}(\mathbb{K}) = \{P \in \mathrm{GL}_{2p}(\mathbb{K}) : PJ_{2p}\,{}^tP = J_{2p}\} \subset \mathrm{GL}_{2p}(\mathbb{K}).$$

Le groupe symplectique mérite le nom de « groupe classique », mais il apparaît moins souvent que les groupes orthogonaux ou unitaires dans le

cursus typique d'un agrégatif. Une première étude de ce groupe (version symplectique du théorème de la base adaptée, connexité, décomposition polaire) fait l'objet de l'exercice D.17.

5. Lien avec la réduction

Nous avons décrit les orbites pour l'action de congruence de $\mathrm{GL}_n(\mathbb{R})$ sur les matrices symétriques et sur les matrices antisymétriques. Il n'est pas beaucoup plus cher de comprendre les orbites pour l'action de $\mathrm{O}(n)$. Celle-ci est particulièrement intéressante à cause de ses liens avec l'action par conjugaison. On appelle naturellement $\mathrm{O}(n)$-congruence (resp. $\mathrm{O}(n)$-similitude) la restriction au groupe orthogonal de l'action de congruence (resp. similitude) de $\mathrm{GL}_n(\mathbb{R})$.

5.1. Lemme. *Deux matrices de $\mathscr{M}_n(\mathbb{R})$ sont $\mathrm{O}(n)$-congruentes si et seulement si elles sont $\mathrm{O}(n)$-semblables.*

Démonstration. Soit $P\in\mathrm{O}(n)$. Comme ${}^tP=P^{-1}$, les relations $A'=PA{}^tP$ et $A'=PAP^{-1}$ sont équivalentes. □

5.2. Remarque. Ce résultat, somme toute très humble, n'en est pas moins lumineux. C'est le sésame permettant de passer du monde des endomorphismes au monde des formes quadratiques. On le voit en action dans la preuve qui suit.

Le résultat est bien connu : toute matrice symétrique réelle est $\mathrm{O}(n)$-semblable à une matrice diagonale réelle ou, dit autrement, qu'une matrice symétrique réelle est diagonalisable en base orthonormée. Mieux encore, en jouant sur cette équivalence entre $\mathrm{O}(n)$-congruence et $\mathrm{O}(n)$-similitude, nous allons obtenir un lien entre la signature d'une matrice symétrique et le signe de ses valeurs propres.

5.3. Proposition. *Toute matrice symétrique réelle est diagonalisable dans une base orthonormée. Si p (resp. q) est égal au nombre de ses valeurs propres strictement positives (resp. strictement négatives), alors (p,q) est sa signature.*

Démonstration. Soit A une matrice symétrique de $\mathscr{S}_n(\mathbb{R})$. Montrons tout d'abord que ses valeurs propres sont réelles.

On peut voir cette matrice comme la matrice d'un endomorphisme φ de $\mathbb{C}^n$ muni de la forme hermitienne canonique $\langle\cdot,\cdot\rangle$, *i.e.* la forme hermitienne

telle que la base canonique (e_i) soit orthonormée. Comme (e_i) est orthonormée, l'adjoint φ^* pour la forme hermitienne admet alors pour matrice la transconjuguée A^* et donc $\varphi^* = \varphi$, *i.e.* φ est autoadjointe, puisque A est symétrique et réelle.

Soit λ une valeur propre de φ et v un vecteur propre pour λ. On a :

$$\lambda\langle v,v\rangle = \langle v,\lambda v\rangle = \langle v,\varphi(v)\rangle = \langle \varphi^*(v),v\rangle = \langle \varphi(v),v\rangle = \langle \lambda v,v\rangle = \overline{\lambda}\langle v,v\rangle.$$

Comme v est non nul, $\langle v,v\rangle$ est aussi non nul, ce qui implique que λ est réel.

Revenons maintenant sur $\mathbb{R}^n$ vu comme espace euclidien muni de la forme $\langle\cdot,\cdot\rangle$ restreinte à $\mathbb{R}^n$ muni de la base orthonormée (e_i). Soit $\varphi_{\mathbb{R}}$ la restriction de φ à $\mathbb{R}^n$, qui a toujours A pour matrice dans la base canonique et qui est bien un endomorphisme, car A est réelle. Soit f_1 un vecteur propre (il en existe puisque les valeurs propres de A sont réelles), que l'on suppose normé. Alors, l'orthogonal de f_1 est stable par $\varphi_{\mathbb{R}}$, car $w \in f_1^{\perp}$ implique

$$\langle f_1, \varphi_{\mathbb{R}}(w)\rangle = \langle \varphi_{\mathbb{R}}(f_1), v\rangle = \lambda\langle f_1, v\rangle = 0.$$

De plus, l'égalité $\langle v,\varphi_{\mathbb{R}}(v')\rangle = \langle \varphi_{\mathbb{R}}(v),v'\rangle$ valable sur $\mathbb{R}^n$ (car vraie sur $\mathbb{C}$) reste valable par restriction sur $f_1^{\perp}$. Donc la restriction de $\varphi_{\mathbb{R}}$ à $f_1^{\perp}$ est encore autoadjointe.

On laisse au lecteur le soin de mettre en place une récurrence sur la dimension pour montrer que tout endomorphisme autoadjoint d'un espace euclidien est diagonalisable en base orthonormée.

Reste à montrer la dernière assertion. Soit $\lambda_1,\cdots,\lambda_p$ ses valeurs propres strictement positives, $\lambda_{p+1},\cdots,\lambda_{p+q}$ ses valeurs propres strictement négatives (les autres étant nulles). D'après le lemme précédent, A est congruente à la matrice diagonale

$$\operatorname{diag}(\lambda_1,\cdots,\lambda_p,\lambda_{p+1},\cdots,\lambda_{p+q},0,\cdots,0).$$

Comme dans la preuve du théorème 1.2, on montre alors que A est congruente à $\mathrm{I}_{(p,q)}$, et donc que A est de signature (p,q). □

Le cas antisymétrique se traite de même.

5.4. Proposition. *Les valeurs propres d'une matrice antisymétrique réelle sont toutes dans $i\mathbb{R}$.*

Démonstration. C'est analogue au cas symétrique, sauf que l'on obtient $\varphi^* = -\varphi$, puis $\overline{\lambda} = -\lambda$. □

6. Classification des coniques

Toute courbe cycloïdale ayant pour base le centre imaginaire du plan de l'infini engendre en tournant autour d'un axe isotrope une quadrique de révolution.

Tous les chemins mènent à Rome, Jean Boyer, 1949.

Dans cette partie, nous discuterons brièvement de la classification affine des coniques. Tout d'abord, nous donnerons une définition des coniques qui ne coïncide pas avec la définition classique et qui en diffère même dramatiquement dès que le corps de base n'est plus algébriquement clos. Cette définition amènera à nous intéresser moins aux zéros de la fonction polynomiale de degré 2 qu'à la fonction elle-même, à scalaire multiplicatif près. Ensuite, nous allons, non pas classer ces coniques sur tout corps (ce qui constitue un problème difficile), mais relier la classification de ces coniques modulo l'action naturelle du groupe affine à celle des formes quadratiques, vues comme fonctions homogènes de degré 2. La classification sur les corps usuels $\mathbb{C}$, $\mathbb{R}$, $\mathbb{F}_q$ en découlera aussitôt.

On travaille sur un corps $\mathbb{K}$ de caractéristique différente de 2. Une conique de l'espace affine $\mathbb{A}^2 = \mathbb{K}^2$ est plutôt vue comme l'ensemble des zéros d'une fonction polynomiale de degré 2 :

$$f(x, y) = \alpha x^2 + \beta xy + \gamma y^2 + \mu x + \nu y + \omega.$$

Toutefois, dans l'objectif d'une classification des coniques en lien avec celle des formes quadratiques, il peut être plus judicieux de définir la conique comme étant la fonction f à un scalaire près.

6.1. Mise en garde. S'il est clair que deux fonctions polynomiales égales à un scalaire près déterminent le même ensemble de zéros, la réciproque est loin d'être vraie. Par exemple sur $\mathbb{R}$, le fonctions $x^2 + y^2 + 1$ et $x^2 + 1$ ont le même ensemble de zéros, c'est-à-dire l'ensemble vide ; pourtant, elles ne sont pas égales à scalaire près. En revanche, si le corps $\mathbb{K}$ est algébriquement clos, le *Nullstellensatz* de Hilbert assure que la réciproque est vraie et donc que les deux définitions (disons la géométrique et l'algébrique) des coniques sont équivalentes.

On opte donc pour la définition algébrique suivante.

6.2. Définition. Une conique est une fonction polynomiale sur $E = \mathbb{K}^2$ de degré 2 (non nécessairement homogène) à scalaire multiplicatif non nul près.

Nous allons nous intéresser à l'espace des fonctions polynomiales de degré 2 muni de l'action du groupe affine.

Pour f comme ci dessus, notons A_f la matrice dans la base canonique :

$$A_f = \begin{pmatrix} \alpha & \beta/2 & \mu/2 \\ \beta/2 & \gamma & \nu/2 \\ \mu/2 & \nu/2 & \omega \end{pmatrix}.$$

Nous allons tout d'abord montrer que la donnée de la fonction f détermine de façon unique la donnée de A_f. (Attention, f n'est pas un polynôme mais une fonction polynomiale !). Ceci est assuré par le lemme suivant.

6.3. Lemme. *Le morphisme d'anneau qui au polynôme P de $\mathbb{K}[X,Y]$ associe la fonction polynôme $(x,y) \mapsto P(x,y)$ sur $\mathbb{K}^2$ est injectif sur le sous-espace $\mathbb{K}[X,Y]_2$ des polynômes de degré inférieur à 2.*

En termes matriciels, si pour tout x, y de $\mathbb{K}^2$, on a :

$$(x \quad y \quad 1)\, A_f \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0,$$

alors $A_f = 0$.

Démonstration. Considérons le morphisme de $\mathbb{K}$-algèbres qui à un polynôme P de $\mathbb{K}[X]$ associe la fonction polynôme $x \mapsto P(x)$ sur $\mathbb{K}$. Si le corps est infini, alors le morphisme est injectif puisqu'un polynôme non nul ne peut avoir qu'un nombre fini de racines. Si $\mathbb{K}$ est fini de cardinal q, alors son noyau est non nul, puisqu'il contient $X^q - X$. Puisqu'un polynôme non nul qui s'annule sur $\mathbb{K}$ doit avoir au moins q racines, il doit être de degré au moins q. Le noyau, qui est un idéal principal, est donc l'idéal engendré par $X^q - X$.
Comme $q > 2$ par hypothèse sur le corps, il vient que sa restriction au sous-espace des polynômes de degré inférieur à 2 est injective.
Supposons maintenant que $P = \alpha X^2 + \beta XY + \gamma Y^2 + \mu X + \nu Y + \omega$ est un polynôme de $\mathbb{K}[X,Y]_2$ tel que $P(x,y) = 0$ pour tout couple (x,y) de $\mathbb{K}^2$. Fixons y dans $\mathbb{K}$, il vient que la fonction $x \mapsto P(x,y)$ est une fonction polynomiale de degré inférieur à 2 nulle sur $\mathbb{K}$. Par ce qui précède, le polynôme $\alpha X^2 + \beta X(y+\mu) + \gamma y^2 + \nu y + \omega$ en X est nul. Ses coefficients α, $\beta(y+\mu)$, $\gamma y^2 + \nu y + \omega$ sont nuls, et ce pour tout y dans $\mathbb{K}$. En utilisant encore une fois le résultat en une indéterminée, il vient que $\alpha = \beta = \gamma = \mu = \nu = \omega = 0$ comme désiré.
La dernière assertion découle immédiatement de ce résultat. □

En particulier, d'après le lemme, toute fonction polynomiale de degré 2 se décompose de façon unique en $f = q+\ell+c$, où q est une forme quadratique non nulle sur $\mathbb{K}^2$, ℓ une forme linéaire de $E^* = (\mathbb{K}^2)^*$ et c une constante de $\mathbb{K}$. L'ensemble $\mathscr{C}$ des coniques, selon notre définition, est donné par

$$\mathscr{C} = (\mathscr{Q}\backslash\{0\} \times E^* \times \mathbb{K})/\mathbb{K}^*,$$

où $\mathscr{Q}$ est l'espace des formes quadratiques, et où le groupe multiplicatif $\mathbb{K}^*$ agit sur $\mathscr{Q}\backslash\{0\} \times E^* \times \mathbb{K}$ par :

$$\lambda \cdot (q, \ell, c) = (\lambda q, \lambda\ell, \lambda c).$$

La classification affine des coniques consiste donc à déterminer les orbites pour l'action du groupe affine $\mathrm{GA}_2(\mathbb{K})$ sur $\mathscr{C}$ donné par :

$$(g \cdot f)(x, y) = f\big(g^{-1}(x, y)\big), \quad g \in \mathrm{GA}_2(\mathbb{K}),\ f \in \mathscr{C},\ (x, y) \in \mathbb{K}^2.$$

On vérifie naturellement que l'action est bien définie, c'est-à-dire qu'elle ne dépend pas du choix du représentant de f modulo la multiplication par λ. Notons que l'on a tordu g par le passage à l'inverse afin d'obtenir une action à gauche.

On note donc $\mathscr{O}_f \subset \mathscr{C}$ l'orbite de f pour cette action et l'on aimerait comprendre la décomposition de $\mathscr{C}$ en orbites.

Pour toute fonction polynomiale $f(x, y) = \alpha x^2 + \beta xy + \gamma y^2 + \mu x + \nu y + \omega$, on note $q_f + \ell_f + c_f$ la décomposition de la fonction f, et Q_f la forme quadratique homogénéisée de f sur $\mathbb{K}^3$, c'est-à-dire :

$$q_f(x, y) = \alpha x^2+\beta xy+\gamma y^2,\ Q_f(x, y, z) = \alpha x^2+\beta xy+\gamma y^2+\mu xz+\nu yz+\omega z^2.$$

Remarquons que Q_f (qui par le lemme précédent ne dépend que de f) est la forme quadratique de matrice A_f dans la base canonique de $\mathbb{K}^3$. Le lemme suivant permet de relier la classification des coniques à celle des formes quadratiques.

6.4. Lemme. *Soit f' une conique de l'orbite $\mathscr{O}_f$, alors la forme quadratique $q_{f'}$ (resp. $Q_{f'}$) est congruente à λq_f (resp. λQ_f) pour un λ de $\mathbb{K}^*$.*

Démonstration. Pour $g \in \mathrm{GA}_2(\mathbb{K})$, on voit facilement à partir de l'égalité (on assimile les points, les vecteurs avec leurs coordonnées dans la base canonique) :

$$\begin{aligned}(g \cdot f)(M) &= f\big(g^{-1}(M)\big) = f\big(\vec{g}^{-1}(\overrightarrow{OM}) + g(O)\big)\\ &= q_f\big(\vec{g}^{-1}(\overrightarrow{OM}) + g(O)\big) + \ell_f\big(\vec{g}^{-1}(\overrightarrow{OM}) + g(O)\big) + c_f\\ &= q_f\big(\vec{g}^{-1}(\overrightarrow{OM})\big) + 2b_f\big(\vec{g}^{-1}(\overrightarrow{OM}), \overrightarrow{Og(O)}\big)\\ &\qquad + \ell_f\big(\vec{g}^{-1}(\overrightarrow{OM})\big) + q_f\big(\overrightarrow{Og(O)}\big) + \ell_f\big(g(O)\big) + c_f,\end{aligned}$$

où b_f est la forme bilinéaire associée à q_f.

On voit donc que la forme quadratique $q_{g\cdot f}$ associée à $g \cdot f$ est égale à $q_f \circ \overrightarrow{g}^{-1}$. Comme par hypothèse $f' = \lambda f \circ g^{-1}$ pour un λ de $\mathbb{K}^*$ et un g de $\mathrm{GA}_2(\mathbb{K})$, on obtient la première assertion.

Pour la seconde assertion, on remarque que si $f'(x,y) = \lambda f\big(g^{-1}(x,y)\big)$, alors

$$(x \quad y \quad 1)\, A_{f'} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \lambda (x,y,1)\, {}^tP A_f P (x \quad y \quad 1),$$

où P désigne la matrice, dans la base canonique de $\mathbb{K}^3$, de l'image de g^{-1} par le morphisme $\mathrm{GA}_2(\mathbb{K}) \to \mathrm{GL}_3(\mathbb{K})$ (voir proposition II-5.4.1).

Par le lemme qui précède, cela implique la relation $A_{f'} = \lambda\, {}^tP A_f P$, d'où la seconde assertion. □

Toutefois, l'orbite de f ne peut pas être comprise à partir de la seule donnée de l'orbite de congruence des formes quadratiques q_f et Q_f. Il est encore nécessaire de définir une action plus adaptée.

Sur les formes quadratiques, l'action de congruence et l'action de $\mathbb{K}^*$ par multiplication commutent, elles peuvent donc s'associer en une seule action de $\mathbb{K}^* \times \mathrm{GL}_n(\mathbb{K})$ donnée par $(\lambda, \overrightarrow{g}) \cdot q = \lambda q \circ \overrightarrow{g}^{-1}$, que l'on appellera action de congruence projective.

Matriciellement, les formes q et q' sont dans la même orbite pour la congruence projective si et seulement si leurs matrices respectives A et A' dans une base fixée vérifient $A' = \lambda P A\, {}^tP$, pour $\lambda \in \mathbb{K}^*$, $P \in \mathrm{GL}_2(\mathbb{K})$.

Pour toute forme quadratique q de $\mathscr{Q}$, on note S_q le stabilisateur de l'orbite (par congruence) de q dans $\mathbb{K}^*$:

$$S_q := \{\lambda \in \mathbb{K}^*,\ \lambda q \text{ congruente à } q\}.$$

6.5. Exemple. Si $\mathbb{K} = \mathbb{C}$, alors il y a deux orbites pour l'action de congruence projective sur les formes quadratiques non nulles. On peut les représenter par $q_1 = x^2$ et $q_2 = x^2 + y^2$. On a $S_{q_1} = S_{q_2} = \mathbb{C}^*$, puisque les classes de congruence sur $\mathbb{C}$ sont classées par le rang et que la multiplication par un scalaire non nul ne change pas le rang.

Si $\mathbb{K} = \mathbb{R}$, alors il y a cette fois-ci trois orbites. On peut les représenter par les formes normales $q_1 = x^2$, $q_2 = x^2 + y^2$, $q_3 = x^2 - y^2$. On voit que $S_{q_1} = S_{q_2} = \mathbb{R}^{+*}$ et $S_{q_3} = \mathbb{R}^*$, en utilisant la classification des formes quadratiques réelles par la signature.

Si $\mathbb{K} = \mathbb{F}_q$, alors on a trois orbites. On peut les représenter par $q_1 = x^2$, $q_2 = x^2 + y^2$, $q_3 = x^2 + \zeta y^2$, où ζ est un élément fixé de $\mathbb{F}_q \backslash \mathbb{F}_q^2$, c'est-à-dire un non carré. Multiplier par un scalaire une forme quadratique en

dimension 2 ne change pas son discriminant, et comme le discriminant est un invariant total, on a $S_{q_2} = S_{q_3} = \mathbb{F}_q^*$ et $S_{q_1} = \mathbb{F}_q^{*2}$.

On a alors la proposition suivante.

6.6. Proposition. *On suppose que $q_i := a_i x^2 + b_i y^2$, $i \in I$ est un système de représentants de formes quadratiques non nulles sur $\mathbb{K}^2$ à congruence projective près; on suppose également que $a_i \neq 0$. Alors, chaque orbite pour l'action du groupe affine sur l'ensemble $\mathscr{C}$ des coniques contient un et un seul élément de la forme suivante :*
- premier cas : $f_{\overline{k}} := a_i x^2 + b_i y^2 + k$, $\overline{k} \in \mathbb{K}/S_{q_i}$, si $b_i \neq 0$;
- second cas : $p_1 := x^2 + y$ ou $e_{\overline{k}} := x^2 + k$, $\overline{k} \in \mathbb{K}/\mathbb{K}^{*2}$, si $(a_i, b_i) = (1, 0)$.

6.7. Remarque. On a vu, proposition B.3.1, que toute orbite de congruence contient une matrice diagonale $\mathrm{diag}(a, b)$. Si la matrice est de rang 1, on peut, quitte à échanger les deux vecteurs de base, supposer que b est nul et donc quitte à diviser par a_i, on peut supposer $a_i = 1$. La disjonction des cas de la proposition n'est donc pas restrictive.

Démonstration. Nous montrons dans un premier temps que toute orbite contient au moins un des représentants proposés. On montrera ensuite que ces fonctions sont dans des orbites distinctes.

Soit $f = q_f + \ell_f + c_f$ un polynôme de degré 2. Alors, par hypothèse, il existe $(\lambda, \overrightarrow{g}) \in \mathbb{K}^* \times \mathrm{GL}_2(\mathbb{K})$ tel que $\lambda q_f\big(\overrightarrow{g}(x, y)\big) = a_i x^2 + b_i y^2$, $i \in I$. On se convainc alors facilement que $\lambda f\big(g(x, y)\big) = a_i x^2 + b_i y^2 + \mu x + \nu y + \omega$, pour g affine d'endomorphisme associé $\overrightarrow{g}$.

Premier cas. Si a_i, b_i sont tous deux non nuls, alors en notant t la translation $t(x, y) = \big(x - \mu/(2a_i), y - \nu/(2b_i)\big)$. On obtient $\lambda f\big(gt(x, y)\big) = a_i x^2 + b_i y^2 + \omega'$.

Maintenant, pour tout $\zeta_i \in S_{q_i}$, il existe un automorphisme g_i tel que $\zeta_i^{-1} q_i = q_i \circ g_i$ et donc $\lambda f\big(gtg_i(x, y)\big) = \zeta_i^{-1}(a_i x^2 + b_i y^2) + \omega'$. On a donc $a_i x^2 + b_i y^2 + \zeta_i \omega' \in \mathscr{O}_f$.

Deuxième cas. Si b_i est nul, alors, comme précédemment, on se ramène, à l'aide de la translation de vecteur $\big(x - \mu/(2a_i), y\big)$ et en multipliant par a_i^{-1}, au cas où $\lambda f(x, y) = x^2 + \nu y + \omega$.

Si $\nu \neq 0$. Alors, en posant $g(x, y) = \big(x, \nu^{-1}(y - \omega)\big)$, on obtient $f \circ g = p_1$.

Si $\nu = 0$. Soit k un représentant de la classe modulo $\mathbb{K}^{*2}$ de ω, de sorte que $\omega = \rho^2 k$. En considérant $g(x, y) = (\rho x, y)$, on montre que f est dans l'orbite de $e_1 = x^2 + k$.

Reste à montrer que les fonctions proposées sont dans des orbites distinctes. Comme une forme quadratique de rang 1 n'est pas congruente à une de

rang 2, le lemme ci-dessus implique que les fonctions du premier cas et celles du deuxième cas sont dans des orbites distinctes. Reste à analyser les cas où les fonctions sont du même type.

Premier cas. Soit $f = a_i x^2 + b_i y^2 + k$ et $f' = a_{i'} x^2 + b_{i'} y^2 + k'$ dans la même orbite et notons $q_i = a_i x^2 + b_i y^2 = q_f$ et $q_{i'} = a_{i'} x^2 + b_{i'} y^2 = q_{f'}$. Alors, par le lemme qui précède, il existe λ et g tels que $q_{i'} = \lambda q_i \circ g$. Comme, par hypothèse, les q_i forment un système de représentants de chaque orbite de congruence projective, on a $i = i'$ et donc $\lambda \in S_{q_i}$. On obtient par identification dans la formule $f' = \lambda f \circ g$ que $g(O) = O$ et que $k' = \lambda k$. Et donc, $k' = k$ modulo S_{q_i}.

Deuxième cas. Il suffit d'utiliser le lemme précédent et calculer les rangs des formes quadratiques associées :

$$\mathrm{rg}(Q_{p_1}) = 3,\ \mathrm{rg}\big(Q_{e_{\bar{k}}}\big) = 2,\ \bar{k} \neq 0,\ \mathrm{rg}(Q_{e_0}) = 1.$$

Le fait que les $e_{\bar{k}}$, $\bar{k} \neq 0$, sont dans des orbites distinctes se voit par un calcul immédiat. □

La classification des formes quadratiques exprimée par le théorème 1.2, sur les corps $\mathbb{C}$, $\mathbb{R}$, $\mathbb{F}_q$ donne alors rapidement la classification affine des coniques, que l'on résume dans le tableau suivant.

On note ζ un élément fixé de $\mathbb{F}_q \backslash \mathbb{F}_q^2$.

Cas où q_f est de rang 2

	$\mathbb{C}$	$\mathbb{R}$		$\mathbb{F}_q$	
q_f	x^2+y^2	x^2+y^2	x^2-y^2	x^2+y^2	$x^2+\zeta y^2$
S_q	$\mathbb{C}^*$	$\mathbb{R}^{+*}$	$\mathbb{R}^*$	$\mathbb{F}_q^*$	$\mathbb{F}_q^*$
f	x^2+y^2+1	x^2+y^2+1 x^2+y^2-1	x^2-y^2+1	x^2+y^2+1	$x^2+\zeta y^2+1$
	x^2+y^2	x^2+y^2	x^2-y^2	x^2+y^2	$x^2+\zeta y^2$

Cas où q_f est de rang 1

$\mathbb{C}$	$\mathbb{R}$	$\mathbb{F}_q$
x^2+y	x^2+y	x^2+y
x^2+1	x^2+1	x^2+1
	x^2-1	$x^2+\zeta$
x^2	x^2	x^2

Sur $\mathbb{R}$, on pourra reconnaître avec bonheur ellipse, hyperbole, parabole.

6.8. Exercice (Classification des quadriques). *Généraliser tous ces résultats à la classification des fonctions polynomiales à n variables de degré inférieur ou égal à 2, sous l'action du groupe affine.*

A. Annexe. Généralités sur les formes bilinéaires

Soit $\mathbb{K}$ un corps de caractéristique différente de 2.

A.1. Formes bilinéaires symétriques ou antisymétriques

A.1.1. Définition. Soit E un espace vectoriel sur $\mathbb{K}$. Une application $b : E \times E \to \mathbb{K}$ est appelée *forme bilinéaire* si, pour tout x de E, les applications $y \mapsto b(x,y)$ et $y \mapsto b(y,x)$ sont des formes linéaires sur E.

On lui associe un morphisme $\psi_b : E \to E^*$ de la façon suivante :

$$\forall (x,y) \in E \times E, \quad \langle \psi_b(x), y \rangle = b(x,y),$$

où E^* désigne le dual de E et $\langle \cdot, \cdot \rangle : E^* \times E \to \mathbb{K}$ désigne ici l'évaluation (d'une forme linéaire en un vecteur).

On appelle *noyau* de b le noyau du morphisme associé ψ_b :

$$\operatorname{Ker} b = \operatorname{Ker} \psi_b = \{x \in E \ , \ \ \forall y \in E, \ b(x,y) = 0\}.$$

On dit que b est *non dégénérée* si son noyau est réduit à $\{0\}$.

Le noyau doit être distingué du *cône isotrope* de b, formé des vecteurs $x \in E$ tels que : $b(x,x) = 0$.

A.1.2. Définition. On appelle *rang* de b le rang du morphisme ψ_b ; c'est donc :

$$\operatorname{rg} b = \dim E - \dim \operatorname{Ker} b.$$

A.1.3. Remarque. Si l'on munit E de l'action naturelle de $\mathrm{GL}(E)$, et son dual E^* de l'action $g \cdot \phi = \phi \circ g^{-1}$, $g \in \mathrm{GL}(E)$, $\phi \in E^*$. Alors, le morphisme ψ_b ne commute (en général) pas à l'action. En revanche, si l'on note O_b le sous-groupe des automorphismes de E qui respectent b, on obtient alors sans mal que ψ_b est un morphisme qui commute à l'action de O_b.

A.1.4. Remarque (Formes symétriques et antisymétriques). Il est bien connu qu'une forme b est dite symétrique, resp. antisymétrique, si $b(x,y) = b(y,x)$, resp. $b(x,y) = -b(y,x)$, pour tout couple de vecteurs (x,y). Une forme bilinéaire s'écrit alors de façon unique comme somme d'une forme symétrique et d'une forme antisymétrique. Pour s'en convaincre, contempler la formule[3] :

$$\forall (x,y) \in E \times E, \quad b(x,y) = \frac{b(x,y)+b(y,x)}{2} + \frac{b(x,y)-b(y,x)}{2}.$$

Dans la suite, nous ne considérerons que des formes symétriques ou antisymétriques sur des espaces de dimension finie. En particulier, dans le cas où b est non dégénérée l'endomorphisme ψ_b ci-dessus est un isomorphisme entre E et son dual E^*.

Il y a une bonne et une mauvaise raison à ça : d'un côté, les groupes qui préservent une forme non dégénérée symétrique ou antisymétrique sont des groupes « presque » simples, au contraire des formes générales ; de l'autre, la question d'une « forme normale » pour les matrices à congruence près, si elles ne sont pas symétriques ou antisymétriques, est trop compliquée (même sur $\mathbb{C}$ ou $\mathbb{R}$; voir, par contraste, la classification des formes quadratiques et des formes symplectiques)[4].

Matrices d'une forme bilinéaire

A.1.5. Définition. Soit b une forme bilinéaire sur un espace E de dimension finie n et $\mathbf{e} = (e_1, \ldots, e_n)$ une base de E. On appelle *matrice de b dans* $\mathbf{e}$ et on note $\mathrm{mat}_{\mathbf{e}}(b)$ la matrice $n \times n$ suivante :

$$\mathrm{mat}_{\mathbf{e}}(b) = (a_{ij})_{1 \leqslant i,j \leqslant n} \quad \text{où} \quad \forall i,j,\ a_{ij} = b(e_i, e_j).$$

A.1.6. Remarque. La matrice de b dans la base $\mathbf{e}$ est également la matrice de l'application linéaire ψ_b dans les bases $\mathbf{e}$ et $\mathbf{e}^*$, où $\mathbf{e}^*$ est la base duale de $\mathbf{e}$. Aussi, le rang de b est également le rang de n'importe laquelle de ses matrices.

A.1.7. Lemme. *Soit b une forme bilinéaire sur E et soit $\mathbf{e}$, $\mathbf{e}'$ deux bases de E. Soit A (resp. A') la matrice de b dans $\mathbf{e}$ (resp. $\mathbf{e}'$) et P la matrice*

[3] La division par 2 dans cette formule est une source inouïe de difficultés et, *in fine*, la principale raison pour laquelle nous excluons de cette étude les corps de caractéristique 2.

[4] Il est d'ailleurs étonnant que ce problème fasse encore l'objet de travaux récents ; voir, par exemple, « Bilinear forms over an algebraic closed field », de B. Corbas et G. D. Williams, au Journal of Pure and Applied Algebra (2001).

de passage de $\mathbf{e}$ *à* $\mathbf{e}'$. *On a :*

$$A' = {}^tPAP \quad (\textit{ou encore : } A = {}^tP^{-1}A'P^{-1}).$$

Démonstration. Soit v et w deux vecteurs de E, X et Y (resp. X' et Y') leurs colonnes de coordonnées dans $\mathbf{e}$ (resp. $\mathbf{e}'$). On a, par définition des matrices P, A et A' :

$$X = PX' \text{ et } Y = PY' \quad b(v,w) = {}^tXAY \text{ et } b(v,w) = {}^tX'A'Y'.$$

On en déduit :

$${}^tX'\,{}^tPAPY' = {}^tX'A'Y'.$$

Comme cela est valable pour X et Y quelconques, le lemme en résulte. □

A.1.8. Corollaire. *La classe du déterminant de la matrice d'une forme bilinéaire modulo les carrés du corps ne dépend pas de la base dans laquelle la matrice est écrite.*

Démonstration. On a en effet, dans les notations du lemme : $\det A = \det(P)^2 \det(A')$. □

A.1.9. Définition (Discriminant). Le *discriminant* d'une forme quadratique non dégénérée q (ou de la forme bilinéaire symétrique associée) sur un espace E de dimension finie est la classe dans $\mathbb{K}^*/(\mathbb{K}^*)^2$ du déterminant de n'importe laquelle des matrices de q.

Quotient par le noyau

Si b est une forme bilinéaire dégénérée sur E, on peut passer au quotient dans $E/\operatorname{Ker} b$ de la façon suivante. Pour x et y éléments de E et x_0, y_0 de $\operatorname{Ker} b$, on a : $b(x+x_0, y+y_0) = b(x,y)$. On définit alors $\bar{b}$ sur $E/\operatorname{Ker} b$ par $\bar{b}(\bar{x},\bar{y}) = b(x,y)$, où $\bar{\ }$ est la classe modulo $\operatorname{Ker} b$. Si $\bar{b}(\bar{x},\bar{y}) = 0$ pour tout y, alors x appartient au noyau de b, *i.e.* $\bar{x} = 0$: ainsi, $\bar{b}$ est non dégénérée.

Le passage de b à $\bar{b}$, ou plutôt de ψ_b à $\psi_{\bar{b}}$, est résumé par le diagramme suivant :

$$\begin{array}{ccc} E & \xrightarrow{\psi_b} & E^* \\ {\scriptstyle\pi}\downarrow & & \uparrow{\scriptstyle {}^t\pi} \\ E/\operatorname{Ker}\psi_b & \xrightarrow[\psi_{\bar{b}}]{\sim} & (E/\operatorname{Ker}\psi_b)^* \end{array}$$

où π est la surjection canonique et ${}^t\pi$ sa transposée, qui est donc injective, car la transposée d'un morphisme surjectif est injective (se rappeler que le noyau de la transposée est égal à l'orthogonal de l'image).

Pour ceux à qui les diagrammes déplaisent, on peut travailler matriciellement, au prix de perdre la naturalité de la construction. On choisit une base $(e_1, \ldots, e_d)$ du noyau de b, que l'on complète en une base $(e_1, \ldots, e_n)$ de E. La matrice de b dans cette base est de la forme :

$$A = \begin{pmatrix} 0 & 0 \\ 0 & \overline{A} \end{pmatrix},$$

où $\overline{A}$ est la matrice de $\bar{b}$ dans la base $(\overline{e_{d+1}}, \ldots, \overline{e_n})$ de $E/\operatorname{Ker} b$.

A.1.10. Définition. On conserve les les notations ci-dessus. On appelle *discriminant réduit* de b et on note $\bar{\delta}(b)$ le discriminant de la forme $\bar{b}$: c'est le discriminant de la matrice $\overline{A}$ (par convention, c'est 0 si b est la forme nulle).

On peut remarquer que $\bar{\delta}(b)$ est la classe modulo les carrés du produit des coefficients non nuls de la matrice de b dans une base orthogonale.

Orthogonalité, espaces isotropes

A.1.11. Définition. Étant donné une forme bilinéaire[5] b sur un espace E, on appelle *orthogonal* d'une partie P et l'on note $P^\perp$ l'ensemble :

$$P^\perp = \{x \in E\ ,\ \ \forall y \in P,\ b(x,y) = 0\}.$$

Si $F \subset E$, son orthogonal $F^\perp$ est un sous-espace vectoriel. Le noyau de la restriction de b au sous-espace F est :

$$\operatorname{Ker} b|_{F\times F} = \{x \in F\ ,\ \ \forall y \in F,\ b(x,y) = 0\} = F \cap F^\perp.$$

A.1.12. Définition. On dit que le sous-espace F est *régulier* si la restriction de b à F est non dégénérée, *i.e.* si l'on a : $F \cap F^\perp = \{0\}$. Dans le cas contraire, on dit que le sous-espace F est *isotrope*. On dit que F est *totalement isotrope* si la restriction de b à F est nulle, *i.e.* si l'on a : $F \subset F^\perp$.

A.1.13. Lemme. *Soit b une forme bilinéaire sur E et F un sous-espace de E. Alors :*

(i) $\dim E + \dim(F \cap \operatorname{Ker} b) = \dim F + \dim F^\perp$;

(ii) *si b est non dégénérée,* $\dim E = \dim F + \dim F^\perp$;

(iii) *si F est régulier, on a :* $E = F \oplus F^\perp$.

[5]On la suppose toujours symétrique ou antisymétrique, ce qui évite de distinguer « orthogonal à gauche » et « orthogonal à droite ».

Démonstration. On commence par prouver (ii). On suppose donc que b est non dégénérée. Considérons l'orthogonal de F au sens de la dualité : $F^\circ = \{\ell \in E^*, \forall y \in F, \ell(y) = 0\}$. D'une part, c'est le noyau de la restriction des formes linéaires, $r_F : E^* \to F^*$. Par le théorème du rang, on a : $\dim F + \dim F^\circ = \dim E$. D'autre part, ψ_b est un isomorphisme et l'on peut constater que $F^\circ = \psi_b(F^\perp)$. En effet, si $\ell = \psi_b(x)$, on a : $\langle \ell, y\rangle = b(x,y)$ pour tout y de E, si bien qu'annuler ℓ équivaut à être orthogonal à x. Par suite, $F^\perp$ et F° ont la même dimension, d'où l'assertion (ii).

Pour b quelconque, on passe au quotient par le noyau de b. Remarquons que $F^\perp$ contient $\operatorname{Ker} b$. Par ailleurs, l'image $\overline{F}$ de F dans $E/\operatorname{Ker} b$ a pour dimension $\dim F - \dim(F \cap \operatorname{Ker} b)$ et l'orthogonal de $\overline{F}$ est l'image de $F^\perp$. L'assertion (ii) appliquée à $E/\operatorname{Ker} b$ donne :

$$\begin{aligned}\dim E - \dim \operatorname{Ker} b &= \dim E/\operatorname{Ker} b = \dim \overline{F} + \dim(F^\perp/\operatorname{Ker} b)\\ &= \dim F - \dim(F \cap \operatorname{Ker} b) + \dim F^\perp - \dim \operatorname{Ker} b,\end{aligned}$$

et l'assertion (i) en découle.

Si l'on suppose enfin que F est régulier, on a : $(F \cap \operatorname{Ker} b) \subset (F \cap F^\perp) = \{0\}$, d'où l'on déduit avec (i) : $\dim F + \dim F^\perp = \dim E$. Ceci entraîne (iii). □

En algèbre linéaire, apparaît assez tôt la notion de transposée d'une application linéaire : si $\varphi : E \to E'$ est une linéaire, sa transposée est l'application linéaire ${}^t\varphi : E'^* \to E^*$, $\ell \mapsto \ell \circ \varphi$. C'est l'unique application linéaire telle que[6] $\langle {}^t\varphi(\ell), y\rangle = \langle \ell, \varphi(y)\rangle$ pour ℓ dans E'^* et $y \in E$.

Or une forme bilinéaire b non dégénérée permet, *via* ψ_b, de faire le va-et-vient entre E et son dual E^*. Vu ce qui précède, il est naturel de transporter la notion de transposée par ψ_b car la matrice de la transposée est la transposée de la matrice.

A.1.14. Lemme. *Soit b une forme bilinéaire non dégénérée sur un espace de dimension finie E.*

(i) *Soit φ un endomorphisme de E. Il existe un unique endomorphisme φ^* tel que :*

$$\forall (x,y) \in E^2, \quad b\big(x, \varphi(y)\big) = b\big(\varphi^*(x), y\big).$$

(ii) *Si φ' est un autre endomorphisme linéaire, on a : $(\varphi \circ \varphi')^* = \varphi'^* \circ \varphi^*$.*

(iii) *Si F est un sous-espace de E stable par φ, alors $F^\perp$ est stable par φ^*.*

[6] La formule fait intervenir deux crochets d'évaluation $\langle \cdot, \cdot \rangle$, l'un est $E'^* \times E' \to \mathbb{K}$, l'autre $E^* \times E \to \mathbb{K}$.

A.1.15. Définition. Avec les notations du lemme, l'endomorphisme φ^* est appelé *adjoint* de φ pour b.

Démonstration. Soit $x \in E$. On cherche $\varphi^*(x)$ de sorte que $b\big(\varphi^*(x), y\big) = b\big(x, \varphi(y)\big)$ pour tout y de E. Mais, on a :

$$\begin{cases} b\big(\varphi^*(x), y\big) = \langle \psi_b\big(\varphi^*(x)\big), y\rangle \\ b\big(x, \varphi(y)\big) = \langle \psi_b(x), \varphi(y)\rangle = \langle {}^t\varphi\big(\psi_b(x)\big), y\rangle. \end{cases}$$

Il est donc nécessaire et suffisant d'avoir : $\psi_b\big(\varphi^*(x)\big) = {}^t\varphi\big(\psi_b(x)\big)$ pour tout x, c'est-à-dire $\psi_b \circ \varphi^* = {}^t\varphi \circ \psi_b$, ce qui, par bijectivité de ψ_b, prouve (i).

L'assertion (ii) vient du fait que la transposition renverse l'ordre des produits alors que la conjugaison les préserve.

Pour l'assertion (iii), on prend x dans $F^\perp$. Pour y de F, on a : $b\big(\varphi^*(x), y\big) = b\big(x, \varphi(y)\big)$, qui est nul car x appartient à $F^\perp$ et $\varphi(y)$ à F. □

On peut résumer ce jeu de miroirs par le diagramme :

$$\begin{array}{ccccc} E & \xrightarrow{\psi_b} & E^* & & E \\ {\scriptstyle\varphi^*}\downarrow & & {\scriptstyle{}^t\varphi}\downarrow & & {\scriptstyle\varphi}\uparrow \\ E & \xrightarrow{\psi_b} & E^* & & E \end{array}$$

Variante matricielle de démonstration. Fixons une base de E, ou, ce qui revient essentiellement au même, supposons que $E = \mathbb{K}^n$, que l'on munit de la base canonique. La forme b est caractérisée par sa matrice B, qui est symétrique ou antisymétrique et inversible ; on a : $b(X, Y) = {}^tXBY$ pour tout $X, Y \in \mathbb{K}^n$. Pour A matrice $n \times n$ quelconque, on cherche A^* telle que ${}^t(A^*X)BY = {}^tXBAY$, *i.e.* ${}^tX\,{}^tA^*BY = {}^tXBAY$. Il est nécessaire et suffisant que ${}^tA^*B = BA$, ou encore : $A^* = B^{-1}\,{}^tAB$, ce qui prouve à nouveau (i).

B. Annexe. Classification des formes quadratiques : théorie classique

B.1. Deux caractérisations

B.1.1. Lemme. *Soit E un espace vectoriel de dimension finie sur un corps de caractéristique différente de 2. Fixons une base $\mathbf{e} = (e_1, \ldots, e_n)$ de E. Soit $q : E \to \mathbb{K}$ une application. Il est équivalent de dire :*

(i) *l'application q est un polynôme homogène de degré 2,* i.e. *il existe des formes linéaires $(\ell_1, \ldots, \ell_p)$ et $(\ell'_1, \ldots, \ell'_p)$ telles que $q = \sum_{k=1}^{p} \ell_k \ell'_k$;*

(ii) *il existe une matrice symétrique* $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ *telle que pour tout* $x = \sum_{i=1}^{n} x_i e_i$, $q(x) = \sum_{i,j} a_{ij} x_i x_j$; *autrement dit :* $q(x) = {}^tXAX$, *où* $X = (x_1, \ldots, x_n) \in \mathbb{K}^n$;

(iii) *il existe une forme bilinéaire symétrique* $b : E \times E \to \mathbb{K}$ *telle que* $q(x) = b(x,x)$ *pour tout* x *de* E.

Dans ce cas, la matrice (a_{ij}) *et la forme bilinéaire* b *sont uniquement déterminées par* q. *Plus précisément, on a :* $4b(x,y) = q(x+y) + q(x-y)$ *pour tout* $(x,y) \in E^2$ *et, d'autre part :* $a_{ij} = b(e_i, e_j)$ *pour tous* i *et* j *et, en particulier,* $a_{ii} = q(e_i)$ *pour tout* i.

B.1.2. Définition. On garde les notations du lemme. Une fonction q satisfaisant aux conditions équivalentes du lemme est appelée *forme quadratique*. La forme bilinéaire symétrique b est appelée *forme polaire* de q. La matrice (a_{ij}) est appelée *matrice de* q (ou de b) dans la base $\mathbf{e}$. Une base pour laquelle (a_{ij}) est diagonale est appelée *base orthogonale* pour q (ou b).

Dans la somme définissant la forme quadratique, les indices i et j ne sont pas ordonnés. Si i est différent de j, le terme en $x_i x_j$ s'ajoute à celui en $x_j x_i$ pour donner, par symétrie, $2a_{ij} x_i x_j$. Par exemple, si l'on part avec $q(x) = x_1^2 + 4x_1x_2 - 5x_2^2$, on a $a_{11} = 1$, $a_{12} = 2$, $a_{21} = 2$, $a_{22} = -5$.

B.1.3. Remarque. (a) La forme q ne détermine pas les formes ℓ_i et ℓ'_i de (i), ni la matrice A. Au contraire, tout le jeu (congruence et changement de base) va consister à faire en sorte d'obtenir des expressions « aussi simples que possible ».

(b) En caractéristique 2, les notions de forme quadratique (q) et de forme bilinéaire symétrique (b) ne coïncident plus. Comme $-1 = 1$, formes bilinéaires symétriques et antisymétriques coïncident, si bien que la correspondance providentielle entre formes quadratiques et formes bilinéaires symétriques n'a plus cours. Voir [33], partie V, pour cette théorie subtile.

Démonstration. Supposons que (i) soit satisfaite et soit $\mathbf{e} = (e_1, \ldots, e_n)$ une base de E. Notons $(e_1^*, \ldots, e_n^*)$ la base duale et écrivons $\ell_k = \sum_{i=1}^{n} \lambda_{ki} e_i^*$ et $\ell'_k = \sum_{j=1}^{n} \lambda'_{kj} e_j^*$ pour k compris entre 1 et p. On a :

$$q = \sum_{k=1}^{p} \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_{ki} \lambda'_{kj} e_i^* e_j^* = \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\sum_{k=1}^{p} \lambda_{ki} \lambda'_{kj} \right) e_i^* e_j^*.$$

Si $a_{ij} = \sum_k (\lambda_{ki}\lambda'_{kj} + \lambda_{kj}\lambda'_{ki})/2$ pour i et j compris entre 1 et n, cela définit une matrice symétrique. En remarquant que $e_i^* e_j^* = e_j^* e_i^*$, et en appliquant q à un vecteur $x = \sum x_i e_i$, pour lequel on a : $x_i = e_i^*(x)$ pour $1 \leqslant i \leqslant n$, on obtient l'expression de l'assertion (ii).

Supposons que q admet, dans une base $\mathbf{e}$, une expression comme dans l'assertion (ii). Pour deux vecteurs $x = \sum_{i=1}^{n} x_i e_i$ et $y = \sum_{j=1}^{b} y_j e_j$ de E, on définit : $b(x,y) = \sum_{i,j} a_{ij} x_i y_j$. C'est manifestement une application bilinéaire symétrique et $b(x,x) = q(x)$.

Supposons qu'il existe une forme bilinéaire b telle que $q(x) = b(x,x)$ pour tout x de E. Fixons une base $(e_1, \ldots, e_n)$ et soit $(e_1^*, \ldots, e_n^*)$ sa base duale. La forme $\ell_i : E \to \mathbb{K}$, $x \mapsto b(x, e_i)$ est linéaire ; par linéarité par rapport à la seconde variable, $b(x,y) = \sum_{i=1}^{n} b(x,e_i) e_i^*(y)$ pour tout (i,j), ce qui donne avec $y = x$: $q = \sum_{i=1}^{n} \ell_i e_i^*$, une expression de la forme (i).

Reste à voir que b et (a_{ij}) sont uniquement déterminées par q. Du constat trivial[7] que $x_i y_i = (x_i + y_i)^2/4 - (x_i - y_i)^2/4$ pour tout x_i et y_i de $\mathbb{K}$, on tire par linéarité : $b(x,y) = q(x+y)/4 - q(x-y)/4$. Aussi b est-elle uniquement déterminée par q, de même que la matrice (a_{ij}) puisque l'on a $a_{ij} = b(e_i, e_j)$ pour tous i,j. □

B.2. Changement de base et congruence

Soit q une forme quadratique sur un espace E de dimension n et soit $\mathbf{e} = (e_1, \ldots, e_n)$ et $\mathbf{e}' = (e_1', \ldots, e_n')$ deux bases de E. On note $P = P_{\mathbf{e},\mathbf{e}'}$ la matrice de passage (ses colonnes sont les coordonnées des vecteurs de $\mathbf{e}'$ dans la base $\mathbf{e}$). Soit A (resp. A') la matrice de q dans $\mathbf{e}$ (resp. $\mathbf{e}'$). Soit x un vecteur de E, X (resp. X') la colonne des coordonnées de x dans $\mathbf{e}$ (resp. $\mathbf{e}'$). On a : $X = PX'$.

D'après le paragraphe précédent, on a deux expressions de $q(x)$:

$$ {}^tX'A'X' = q(x) = {}^tXAX = {}^t(PX')APX' = {}^tX'({}^tPAP)X'. $$

Puisque x (donc X) est quelconque et que la matrice de q dans $\mathbf{e}'$ est déterminée par q, on a :

$$ A' = {}^tPAP. $$

B.2.1. Remarque. La formule $P \cdot A = {}^tPAP$ définit une action *à droite* de $\mathrm{GL}_n(\mathbb{K})$ sur l'espace des matrices symétriques $n \times n$. Pour y remédier, deux façons : définir, pour $g \in \mathrm{GL}_n(\mathbb{K})$ et $A \in \mathscr{S}_n(\mathbb{K})$,

$$ g \cdot A = {}^tg^{-1} A g^{-1} \quad \text{ou} \quad g \cdot A = gA\,{}^tg. $$

La première formule a l'inconvénient d'être compliquée par la présence des inverses, mais c'est parfois un avantage : si, par exemple, l'objet central

[7] Une dernière fois : la caractéristique est différente de 2, c'est ce qui rend les choses simples.

de l'étude est $\mathbb{K}^n$ et que $\mathscr{S}_n(\mathbb{K})$ est interprété comme (les matrices associées à) des fonctions sur $\mathbb{K}$, il est très adapté d'avoir un inverse, voir la remarque I-A.1.13. En revanche, si les matrices symétriques constituent le centre d'intérêt, autant opter pour la simplicité avec le deuxième formule. En tout état de cause, les deux actions sont reliées par précomposition de l'automorphisme (extérieur) $g \mapsto {}^t g^{-1}$ de $\mathrm{GL}_n(\mathbb{K})$ (exercice I-C.4).

Il est bien connu, par exemple par le lemme I-A.1.23, que le noyau de l'action par conjugaison de $\mathrm{GL}_n(\mathbb{K})$ sur $\mathscr{M}_n(\mathbb{K})$ est réduit aux homothéties. Il en va de même de l'action par congruence $\mathrm{GL}_n(\mathbb{K})$ sur $\mathscr{M}_n(\mathbb{K})$ (et même sur le sous-espace des matrices symétriques), mais la preuve est un peu plus torturée.

B.2.2. Proposition. *Soit n un entier strictement positif.*

(i) *Soit A dans $\mathrm{GL}_n(\mathbb{K})$. On suppose que $AS\,{}^t A = S$ pour toute matrice symétrique S de $\mathscr{M}_n(\mathbb{K})$. Alors, $A = \pm \mathrm{I}_n$.*

(ii) *Soit H dans $\mathscr{M}_n(\mathbb{K})$. On suppose que $HS + S\,{}^t H = 0$ pour toute matrice symétrique S de $\mathscr{M}_n(\mathbb{K})$. Alors, $H = 0$.*

Démonstration. (i) On a en particulier que $AS\,{}^t A = S$ pour toute matrice inversible symétrique S et donc $AS^{-1}\,{}^t A = S^{-1}$, car S^{-1} est aussi symétrique. Il vient donc : $S^{-1}AS = {}^t A^{-1}$ et $S^{-1}\,{}^t A^{-1} S = A$ et, finalement : $S^{-1}(A + {}^t A^{-1})S = {}^t A^{-1} + A$. Comme toute droite D peut être réalisée comme droite propre d'une matrice symétrique inversible, il vient comme dans le lemme I-A.1.23 que $A + A^{*-1}$ est une homothétie. On a par conséquent $A\mathrm{I}_n\,{}^t A = \mathrm{I}_n$ (en posant $S = \mathrm{I}_n$) et $A + {}^t A^{-1} = \lambda \mathrm{I}_n$. Cela implique que $2A$ est une homothétie et l'on en déduit en caractéristique différente de 2 que A est également une homothétie. La conclusion $A = \pm \mathrm{I}_n$ est claire.

L'assertion (ii) est analogue. On montre que $S^{-1}(H - {}^t H)S = H - {}^t H$ pour toute matrice symétrique inversible S et donc $H - {}^t H$ est une homothétie. Et l'on conclut, car H est anti-hermitienne (en posant $S = \mathrm{I}_n$ dans l'égalité de départ). □

B.3. Théorèmes de classification (nombres complexes, réels et corps finis)

Le premier ingrédient de la classification classique est, encore une fois, la méthode du pivot de Gauss. Dans les écrits de ce dernier, c'est d'ailleurs pour les formes quadratiques qu'elle apparaît pour la première fois[8] à la fin de la *Théorie du Mouvement des Corps célestes* de 1809.

[8] Voir par exemple l'article [9] de Jean-Luc Chabert. Il faut noter que la méthode du pivot adaptée à la résolution des systèmes linéaires et problèmes afférents était déjà connue au premier siècle en Chine, sous le nom de *Fang cheng* et sous la forme de

B.3.1. Proposition (Existence de bases orthogonales). *Soit $\mathbb{K}$ un corps quelconque et q une forme quadratique sur un espace E de dimension finie. Alors, E admet une base orthogonale pour q.*

Conceptuelle. On procède par récurrence sur la dimension n de l'espace E. Pour $n = 1$, la propriété est triviale. Soit n un entier supérieur à 2, on suppose la propriété vraie pour tout espace de dimension au plus $n-1$. Soit b une forme quadratique sur un espace de dimension n. Si q est nulle, toute base est orthogonale. Sinon, soit e_1 un vecteur non isotrope. La droite $\mathbb{K}e_1$ est régulière donc E est la somme directe de $\mathbb{K}e_1$ et de $e_1^\perp$. Par hypothèse de récurrence, on trouve une base orthogonale de la restriction de q à $e_1^\perp$. En y adjoignant e_1, on obtient une base de E orthogonale pour q, ce qui permet de conclure la récurrence. □

Algorithmique, ex nihilo. On procède par récurrence sur la dimension n de l'espace E. Pour $n = 0$, rien à dire. Pour $n = 1$, toute base est trivialement orthogonale. Soit n entier, on suppose l'assertion de la proposition vérifiée pour tout espace de dimension inférieure à $n - 1$. Soit q une forme quadratique sur E de dimension n. Soit $\mathbf{e} = (e_1, \dots, e_n)$ une base de E et (a_{ij}) la matrice de q dans $\mathbf{e}$. Pour (x_i) élément de $\mathbb{K}^n$ et $x = \sum_{i=1}^n x_i e_i$, on a :

$$q(x) = \sum_{i=1}^{x} a_{ii} x_i^2 + \sum_{1 \leqslant i < j \leqslant n} 2a_{ij} x_i x_j.$$

Si e_1 appartient au noyau de q, *i.e.* si $a_{1j} = 0$ pour tout j, c'est facile : par hypothèse de récurrence, on peut trouver une base $(f_2, \dots, f_n)$ de la somme directe $\mathbb{K}e_2 \oplus \cdots \oplus \mathbb{K}e_n$ qui est orthogonale pour la restriction de q et alors $(e_1, f_2, \dots, f_n)$ convient.

Supposons que a_{11} n'est pas nul. On écrit :

$$q(x) = a_{11}\Big(x_1 + \sum_{j=2}^{n} \frac{a_{1j}}{a_{11}} x_j\Big)^2 + q'(x),$$

où q' est une forme quadratique dans laquelle x_1 n'apparaît pas, *i.e.* dont le noyau contient e_1. Par hypothèse de récurrence, on peut trouver une base $(f_2, \dots, f_n)$ de $\mathbb{K}e_2 \oplus \cdots \oplus \mathbb{K}e_n$ orthogonale pour la restriction de q'. Soit $(e_1^*, f_2^*, \dots, f_n^*)$ la base duale de $(e_1, f_2, \dots, f_n)$. On a alors :

$$q(x) = a_{11}\Big(x_1 + \sum_{j=2}^{n} \frac{a_{1j}}{a_{11}} x_j\Big)^2 + \sum_{j=2}^{n} \alpha_j f_j^*(x)^2,$$

dix-huit exercices dans les *Neuf chapitres sur l'art mathématique*, bien que le fin mot de la théorie des systèmes linéaires n'ait été prononcé que dans la deuxième moitié du dix-neuvième siècle.

En remplaçant e_1 par la somme de e_1 et d'une combinaison linéaire convenable des f_j, $j \geqslant 2$, on obtient une base de E orthogonale pour q.

Supposons que a_{11} soit nul. Alors, un des a_{1j} n'est pas nul. Quitte à permuter deux vecteurs de base, on peut supposer que c'est a_{12}. On rassemble tous les termes où interviennent x_1 et x_2 comme suit :

$$q(x) = 2a_{12}x_1x_2 + x_1\ell_1(x) + x_2\ell_2(x) + q'(x),$$

où les formes linéaires ℓ_1, ℓ_2 et la forme quadratique q' ne font intervenir que $x_3, \ldots, x_n$. Soit $b = 2a_{12}$. On peut alors factoriser :

$$q(x) = b\big(x_1 + \frac{1}{b}\ell_2(x)\big)\big(x_2 + \frac{1}{b}\ell_1(x)\big) - \frac{1}{b}\ell_1(x)\ell_2(x) + q'(x),$$

puis, avec $4uv = (u+v)^2 - (u-v)^2$ et en posant $q'' = -\ell_1\ell_2/b + q'$, on obtient pour $q(x)$:

$$\frac{b}{4}\Big(x_1 + x_2 + \frac{\ell_1(x)}{b} + \frac{\ell_2(x)}{b}\Big)^2 - \frac{b}{4}\Big(x_1 - x_2 - \frac{\ell_1(x)}{b} + \frac{\ell_2(x)}{b}\Big)^2 + q''(x).$$

On a ainsi exprimé la forme quadratique q comme somme des carrés des formes linéaires $f_1^* = e_1^* + e_2^* + \ell_1/b + \ell_2/b$ et $f_2^* = e_1^* - e_2^* - \ell_1/b + \ell_2/b$ et d'une forme quadratique q'' en les $f_j^* = e_j^*$ $(j \geqslant 3)$. Grâce à l'hypothèse de récurrence, on exprime q'' comme une somme pondérée de carrés de combinaisons linéaires de $f_3^*, \ldots, f_n^*$, ce qui permet de conclure la récurrence puisque $(f_i^*)_{1 \leqslant i \leqslant n}$ est une base du dual. □

B.3.2. Remarque. À ce stade, nous avons, partant d'une matrice symétrique A, trouvé une matrice diagonale congruente à A. Dans le cadre de la réduction, c'est-à-dire si l'on remplace « congruente » par « semblable », on est arrivé au bout du problème, puisque la diagonale principale, vue dans $\mathbb{K}^n/\mathfrak{S}_n$, est un invariant total, ou dit autrement, puisque deux matrices diagonales sont semblables si et seulement si elles sont égales à permutation près de la diagonale.

Dans le cas de la congruence, le « si » reste valable mais le « seulement si » ne l'est plus et du coup, le problème est loin d'être achevé. En même temps, ça paraissait trop simple !

Voici un premier exemple immédiat (et déjà largement utilisé) qui invalide le « seulement si ». Il se situe en dimension 1 : la matrice (a) est congruente à la matrice $(\lambda^2 a)$ pour tout a et λ dans $\mathbb{K}$. On pourrait se demander si, essentiellement, il ne s'agirait pas de la seule obstruction, c'est-à-dire que deux matrices diagonales seraient congruentes si et seulement si elles sont égales modulo multiplication par des carrés et à permutation près. L'exemple qui suit en dimension 2 donne de quoi prouver qu'il n'en est

rien :

$$\begin{pmatrix} 1 & 1 \\ b & -a \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} 1 & b \\ 1 & -a \end{pmatrix} = \begin{pmatrix} a+b & 0 \\ 0 & (a+b)ab \end{pmatrix}.$$

Ces choses peuvent se dire très élégamment en termes de présentation du groupe de Witt. Pour en savoir plus sur cette histoire (hédoniste), vous êtes invité à lire [33].

B.3.3. Proposition (Formes quadratiques sur $\mathbb{C}$). *Soit b une forme quadratique sur un espace vectoriel* complexe *de dimension n et soit r son rang. Il existe une base $\mathbf{e} = (e_1, \dots, e_n)$ de E telle que :*

$$q(x) = \sum_{i=1}^{r} x_i^2, \quad \textit{pour tout vecteur } x = \sum_{i=1}^{n} x_i e_i.$$

Le théorème de Sylvester qui vient doit être compris comme un résultat de classification des formes quadratiques, à congruence près, par la signature (et la dimension de l'espace). On trouve même une forme normale particulièrement simple dans chaque classe de congruence.

B.3.4. Théorème (Théorème d'inertie de Sylvester ; version 1**).** *Soit q une forme quadratique sur un espace vectoriel* réel *de dimension n. Il existe deux entiers p et r, $0 \leqslant p \leqslant r$ et une base $\mathbf{e} = (e_1, \dots, e_n)$ de E tels que :*

$$q(x) = \sum_{i=1}^{p} x_i^2 - \sum_{j=1}^{r-p} x_j^2, \quad \textit{pour tout vecteur } x = \sum_{i=1}^{n} x_i e_i.$$

De plus, les entiers p et r ne dépendent que de q :

$$r = \operatorname{rg} q, \quad p = \max\{\dim F,\ F \textit{ s.e.v. t.q. } q|_F \textit{ est définie positive}\}.$$

B.3.5. Définition. On garde les notations du théorème. On appelle *signature* de la forme quadratique q (ou de la forme bilinéaire associée) le couple $(p, r-p)$. On dit parfois que p est le nombre de carrés positifs de q et $r-p$ le nombre de carrés négatifs.

De façon équivalente, on peut l'exprimer en termes matriciels.

B.3.6. Théorème (Théorème d'inertie de Sylvester ; version 2**).** *Soit n un entier naturel et A une matrice symétrique $n \times n$ réelle. Il existe deux entiers p et r, avec $1 \leqslant p \leqslant r$, et une matrice inversible P telle que l'on ait, par blocs :*

$${}^tPAP = \begin{pmatrix} \mathrm{I}_p & 0 & 0 \\ 0 & -\mathrm{I}_{r-p} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

De plus, les entiers p et r ne dépendent que de A et pas de P.

Pour n entier, introduisons l'action du groupe linéaire $\mathrm{GL}_n(\mathbb{R})$ par congruence sur l'espace $\mathscr{S}_n(\mathbb{R})$ des matrices symétriques réelles $n \times n$: pour $g \in \mathrm{GL}_n(\mathbb{R})$ et $A \in \mathscr{S}_n(\mathbb{R})$, on fait agir g sur A par :

$$g \cdot A = gA^t g.$$

B.3.7. Théorème (Théorème d'inertie de Sylvester ; version 3). *Soit n un entier. Soit $A \in \mathscr{S}_n(\mathbb{R})$. Il existe deux entiers p et r, avec $1 \leqslant p \leqslant r$, uniquement déterminés par A, et $g \in \mathrm{GL}_n(\mathbb{R})$ tels que*

$$g \cdot A = \begin{pmatrix} \mathrm{I}_p & 0 & 0 \\ 0 & -\mathrm{I}_{r-p} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Cela signifie que les orbites sont classées par les couples $(p, r-p)$ et qu'une forme normale dans l'orbite correspondant à $(p, r-p)$ est la matrice $\mathrm{I}_{p,r-p}$ qui apparaît dans les versions 2 et 3.

De façon équivalente, pour r et p entiers avec $0 \leqslant p \leqslant r$, soit $\mathscr{O}_{p,r-p}$ l'ensemble des matrices dont la forme quadratique associée possède p carrés positifs et $r-p$ carrés négatifs. La matrice $\mathrm{I}_{p,r-p}$ en est un élément.

B.3.8. Théorème (Théorème d'inertie de Sylvester ; version 4). *Soit n un entier naturel. Les orbites de l'action de $\mathrm{GL}_n(\mathbb{R})$ par congruence sur $\mathscr{S}_n(\mathbb{R})$ sont les $\mathscr{O}_{p,r-p}$, $0 \leqslant p \leqslant r \leqslant n$.*

C. Développement. Une preuve de la loi de réciprocité quadratique... par les formes quadratiques

Voici une application des invariants de congruence à un splendide théorème d'arithmétique auquel sont attachés de grands noms tels que Fermat, Legendre et Euler, qui l'ont conjecturée. Gauss put enfin en proposer les premières preuves[9] complètes. On trouve deux démonstrations classiques à la fin du chapitre I de [34], § 3 et appendice.

[9] Gauss en donne deux dans ses célèbres *Disquisitiones Arithmeticae* (1801), alors qu'il n'a que vingt-quatre ans. Franz Lemmermeyer, auteur du livre *Reciprocity laws* [20], lui en attribue huit en tout, ce qui en fait le co-recordman avec Kronecker, et en recense deux cent trente-trois sur son site `www.rzuser.uni-heidelberg.de/~hb3/chrono.html`.

Symbole de Legendre

On a déjà vu l'importance, dans l'étude des formes quadratiques, de savoir si certains éléments, par exemple le discriminant, sont des carrés. Sur les complexes, la question est claire : tous les éléments sont des carrés. Sur les réels, c'est facile : les carrés sont les réels positifs. Le critère suivant permet de détecter les carrés dans $\mathbb{F}_p$, où p est premier.

C.1. Proposition. *Soit p un nombre premier impair et a un élément de $\mathbb{F}_p^*$. Alors :*

$$a^{\frac{p-1}{2}} = \begin{cases} 1 & \text{si } a \text{ est un carré dans } \mathbb{F}_p^*, \\ -1 & \text{si } a \text{ n'est pas un carré dans } \mathbb{F}_p^*. \end{cases}$$

Démonstration. Considérons les morphismes de groupes $\chi : \mathbb{F}_p^* \to \mathbb{F}_p^*$ et $\lambda : \mathbb{F}_p^* \to \mathbb{F}_p^*$, donnés respectivement par $x \mapsto x^2$ et $x \mapsto x^{(p-1)/2}$. D'après[10] le théorème de Lagrange ou le petit théorème de Fermat, le morphisme composé $\lambda \circ \chi = \chi \circ \lambda : x \mapsto x^{p-1}$ est trivial. En particulier, $\lambda(a) = a^{(p-1)/2}$ admet pour carré 1 donc, par intégrité de $\mathbb{F}_p$, vaut ± 1.

On a une suite de groupes :

$$1 \longrightarrow \{-1, 1\} \longrightarrow \mathbb{F}_p^* \xrightarrow{\chi} \mathbb{F}_p^* \xrightarrow{\lambda} \{-1, 1\} \longrightarrow 1.$$

Vu que $\lambda(a)$ vaut -1 ou 1, il s'agit de montrer que a appartient à $\operatorname{Im}\chi$ si et seulement si a appartient à $\operatorname{Ker}\lambda$. Autrement dit, que $\operatorname{Im}\chi = \operatorname{Ker}\lambda$, ou encore que la suite est exacte. Mais l'on a déjà vu l'inclusion : $\operatorname{Im}\chi \subset \operatorname{Ker}\lambda$. De plus, on a : $|\operatorname{Im}\chi| = |\mathbb{F}^*|/|\operatorname{Ker}\chi| = (p-1)/2$ et $|\operatorname{Ker}\lambda| \leqslant (p-1)/2$ puisque $\operatorname{Ker}\lambda$ est l'ensemble des racines du polynôme $x^{(p-1)/2} - 1$ sur un corps (commutatif). On peut conclure l'égalité $\operatorname{Im}\chi = \operatorname{Ker}\lambda$. □

C.2. Définition. Pour p nombre premier impair et a élément de $\mathbb{F}_p$, on définit le *symbole de Legendre de a* par :

$$\left(\frac{a}{p}\right) = a^{\frac{p-1}{2}} = \begin{cases} 1 & \text{si } a \text{ est un carré dans } \mathbb{F}_p^*, \\ -1 & \text{si } a \text{ n'est pas un carré dans } \mathbb{F}_p^*. \\ 0 & \text{si } a = 0. \end{cases}$$

Il est intéressant de remarquer que le symbole de Legendre peut vivre indifféremment dans tout corps fini ou dans le corps des complexes. Mais dans le cas des corps finis, nous venons de voir qu'il est multiplicatif. En particulier, ce qui peut paraître étonnant, c'est que le produit de deux non carrés est un carré. Notons que, si cette propriété est encore valable pour les complexes (il n'y a pas de non carrés !) et pour les réels (moins fois moins égale plus !), ce n'est pas le cas par exemple pour le corps des rationnels.

[10] Que l'on nous pardonne cet abandon au *name dropping*.

Pour $a \neq 0$, le cas par cas de la proposition ci-dessus peut se reformuler en une formule plus synthétique.

C.3. Corollaire. *Soit p premier impair et a un élément de $\mathbb{F}_p^*$. On a :*

$$\left|\{x \in \mathbb{F}_p \ , \ ax^2 = 1\}\right| = 1 + \left(\frac{a}{p}\right).$$

C.4. Remarque (Calculs pratiques du symbole de Legendre). La multiplicativité du symbole de Legendre (le fait que λ soit un morphisme) permet un calcul efficace lorsque l'on dispose d'une factorisation.

Pour n entier non divisible par p, on souhaite déterminer si n est un carré modulo p. En pratique (dans un exercice, c'est-à-dire en théorie...), si l'on dispose d'une factorisation de n, soit $n = \pm\prod_{i=1}^{r} q_i^{\alpha_i}$ où les q_i sont des nombres premiers distincts, la multiplicativité du symbole de Legendre permet de calculer le symbole de Legendre par :

$$\left(\frac{n}{p}\right) = (\pm 1)^{\frac{p-1}{2}} \prod_{i=1}^{r} \left(\frac{q_i}{p}\right)^{\alpha_i}.$$

En pratique (avec un ordinateur), le coût de la factorisation est prohibitif et l'on préférera calculer directement le reste de la division de n par p (en $O(\log n \log p)$ opérations) et sa puissance $(p-1)/2$-ème par un algorithme d'exponentiation rapide (en $O(\log^3 p)$ opérations, voir [10], p. 34). Mais l'on peut améliorer cette complexité en généralisant le symbole de Legendre, grâce à la loi de réciprocité quadratique.

Aparté. On peut utiliser le symbole de Legendre pour un algorithme semi-naïf[11] de factorisation. Soit n entier naturel. On cherche à écrire $n = x^2 - y^2$, ce qui donnera deux facteurs de n, en l'occurrence $x - y$ et $x + y$. Pour cela, on fait décroître y à partir de $\sqrt{n}$ et l'on cherche si $n + y^2$ est un carré entier. Mais, calculer une racine carrée pour voir si c'est un entier, c'est une opération coûteuse.

Au lieu de cela, on regarde si $n+y^2$ est un carré modulo les entiers 11, 63, 64, 65. Si ce n'est pas le cas pour l'un d'entre eux, on est sûr que $n + y^2$ n'est pas un carré entier. Si $n + y^2$ est un carré modulo 11, 63, 64, 65, on calcule sa racine carrée. Si la racine carrée est un entier, on a gagné. Sinon, on a fait le calcul pour rien : mais l'on peut vérifier (et c'est là que le choix de ces entiers se révèle judicieux) que moins de 1 % des classes modulo $m := \text{ppcm}(11, 63, 64, 65)$ ne sont pas des carrés modulo m : aussi, on gagne beaucoup en efficacité. Voir [10], § 1.7.2.

[11] C'est-à-dire plus efficace que la méthode naïve consistant à diviser par les entiers successifs.

Loi de réciprocité quadratique

C.5. Théorème. *Soit p et q deux nombres premiers impairs distincts. Alors :*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}.$$

Démonstration. L'idée est de calculer de deux façons différentes le cardinal modulo p de la « sphère » définie sur $\mathbb{F}_q$:

$$X = \Big\{(x_1, \dots, x_p) \in \mathbb{F}_q^p \ , \ \sum_{i=1}^{p} x_i^2 = 1\Big\}.$$

D'une part, on fait agir le groupe cyclique $\mathbb{Z}/p\mathbb{Z}$ sur X ; d'autre part, on remplace la forme quadratique $\sum x_i^2$ par une forme qui lui est congruente, pour laquelle le cardinal de X est facile.

D'une part, on a :

$$(\spadesuit) \qquad |X| = \left(\frac{p}{q}\right) + 1 \mod p.$$

En effet, faisons agir $\mathbb{Z}/p\mathbb{Z}$ par permutation cyclique sur $\mathbb{F}_q^p$:

$$\forall k \in \mathbb{Z}/p\mathbb{Z},\ \forall (x_1, \dots, x_p) \in \mathbb{F}_q^p, \quad k \cdot (x_1, \dots, x_n) = (x_{1+k}, \dots, x_{n+k}),$$

où les indices sont vus modulo p : $x_{\ell+p} = x_\ell$ pour tout ℓ.

Les orbites de $\mathbb{Z}/p\mathbb{Z}$ dans X sont de deux types : d'un côté, les singletons $\{(x, \dots, x)\}$, où $x \in \mathbb{F}_q$, ont un stabilisateur trivial ; toutes les autres orbites ont un stabilisateur non trivial, nécessairement égal à $\mathbb{Z}/p\mathbb{Z}$ entier. Le nombre d'orbites à stabilisateur trivial est le nombre de solutions x à l'équation : $px^2 = 1$ dans $\mathbb{F}_q$, donné par le corollaire C.3 ; le cardinal des autres orbites est divisible par p. La formule ($\spadesuit$) en résulte.

D'autre part, on a :

$$(\heartsuit) \qquad |X| = \Big(q^d + (-1)^{\frac{p-1}{2}\frac{q-1}{2}}\Big) q^d, \quad \text{où } d = \frac{p-1}{2}.$$

En effet, les matrices $p \times p$ à coefficients dans $\mathbb{F}_q$

$$\mathrm{I}_p = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \quad \text{et} \quad A = \begin{pmatrix} 0 & 1 & & & & & & \\ 1 & 0 & & & & & & \\ & & 0 & 1 & & & & \\ & & 1 & 0 & & & & \\ & & & & \ddots & & & \\ & & & & & 0 & 1 & \\ & & & & & 1 & 0 & \\ & & & & & & & a \end{pmatrix},$$

où $a = (-1)^{(p-1)/2} = (-1)^d$, sont symétriques et ont même rang p et même déterminant 1 (donc même discriminant) : sur $\mathbb{F}_p$, cela signifie qu'elles sont congruentes ! Dit autrement, un changement linéaire de variables identifie X à l'ensemble

$$X' = \big\{(y_1, \dots, y_d, z_1, \dots, z_d, t) \in \mathbb{F}_q^p \ , \ \ 2\,(y_1 z_1 + \cdots + y_d z_d) + at^2 = 1\big\}.$$

Distinguons deux types de points $(y_1, \dots, y_d, z_1, \dots, z_d, t)$ de X' :
- les points tels que $y_1 = \cdots = y_d = 0$: chaque valeur de t telle que $at^2 = 1$ détermine q^d tels points, ce qui par le corollaire C.3 donne $q^d(1+a^{(q-1)/2})$ points en tout ;
- les points pour lesquels au moins un des y_i n'est pas nul : une fois fixés $(y_1, \dots, y_d)$ et t, il reste à choisir un élément $(z_1, \dots, z_d)$ dans un hyperplan affine de $\mathbb{F}_q^d$; il y a $(q^d - 1)qq^{d-1}$ tels points en tout.

En faisant la somme des nombres de points de chaque type, on obtient la formule ($\heartsuit$).

En comparant les deux expressions de $|X|$, il vient :

$$|X| = \Big(q^{\frac{p-1}{2}} + (-1)^{\frac{p-1}{2}\frac{q-1}{2}}\Big) q^{\frac{p-1}{2}} = \Big(\frac{p}{q}\Big) + 1 \mod p$$

et donc, en termes de symboles de Legendre :

$$|X| = \Big[\Big(\frac{q}{p}\Big) + (-1)^{\frac{p-1}{2}\frac{q-1}{2}}\Big]\Big(\frac{q}{p}\Big) = \Big(\frac{p}{q}\Big) + 1 \mod p,$$

ce qui donne, après simplification, la loi de réciprocité quadratique. □

C.6. Remarque (Symbole de Kronecker-Jacobi et calcul pratique du symbole de Legendre). Le symbole de Legendre est multiplicatif par rapport à la variable du haut ; on peut forcer la multiplicativité par rapport à la variable du bas en définissant, pour a et b entiers quelconques, leur symbole de Kronecker-Jacobi $\left(\frac{a}{b}\right)$: en gros, si $b = \prod p_i$ où les p_i sont des nombres premiers, c'est $\prod \left(\frac{a}{p_i}\right)$. On peut alors prouver la loi de réciprocité quadratique C.5 pour p et q impairs non nécessairement premiers. On obtient alors une méthode de calcul du symbole de Jacobi-Kronecker $\left(\frac{a}{b}\right)$ en $O(\log^2 N)$ opérations, où $N = \max(|a|, |b|)$. Voir [10], § 1.4.2.

C.7. Remarque (culturelle). En première approche, la loi de réciprocité quadratique étonne, car elle établit un lien entre le nombre de solutions de deux équations sur des corps différents, l'équation $x^2 = p$ dans $\mathbb{F}_q$ et l'équation $x^2 = q$ dans $\mathbb{F}_p$. Si l'on va plus loin, elle devient la porte d'entrée vers un monde mathématique remarquablement vaste et profond, celui de

la *correspondance de Langlands*, dont relève la preuve du « grand théorème de Fermat » par Wiles en 1994. Dans la rubrique *people*[12], c'est pour leurs contributions à cette théorie que Vladimir Drinfel'd, Laurent Lafforgue et Ngô Bao Chau ont reçu la médaille Fields. Pour une introduction accessible à ce sujet, voir l'annexe F du livre de Pierre Colmez [11].

D. Exercices

Rang et signature

D.1. Exercice. *Soit q une forme quadratique réelle définie (on dit aussi anisotrope), c'est-à-dire telle que l'équation $q(x) = 0$ ait pour seule solution : $x = 0$. Montrer qu'elle est soit définie positive, soit définie négative par deux méthodes, l'une avec les invariants de congruence, l'autre en utilisant le théorème des valeurs intermédiaires.*

D.2. Exercice

1. *Soit f, g deux formes linéaires indépendantes sur un espace vectoriel E de dimension finie sur un corps $\mathbb{K}$ de caractéristique différente de 2. On définit sur E la forme quadratique $q(u) = f(u)g(u)$. Quel est le rang de q ? Sur $\mathbb{R}$, montrer que la signature de q est $(1,1)$.*

   On peut par exemple réaliser f et g comme des formes coordonnées en prolongeant (f,g) en une base de E^* et en prenant l'antéduale.

2. *Soit λ_i, $1 \leqslant i \leqslant 3$, trois scalaires non nuls. Montrer que le polynôme $\lambda_1 X_1^2 + \lambda_2 X_2^2 + \lambda_3 X_3^2$ est irréductible sur $\mathbb{K}[X_1, X_2, X_3]$.*

   On montrera que s'il se factorise, c'est en facteurs homogènes de degré 1.

D.3. Exercice (Formes quadratiques liées à la trace)

1. *Montrer que $A \mapsto \operatorname{tr}(A^2)$ définit une forme quadratique sur l'espace vectoriel $\mathcal{M}_n(\mathbb{R})$.*

2. *Étudier sa restriction sur le sous-espace des matrices symétriques, puis sur le sous-espace des matrices antisymétriques. Montrer que ces deux sous-espaces sont orthogonaux.*

   On note que $(A,B) \mapsto \operatorname{tr}(AB)$ est la forme bilinéaire symétrique associée. On rappelle que $\operatorname{tr}(A) = \operatorname{tr}({}^tA)$.

3. *En déduire sa signature.*

   Que peut-on dire des valeurs propres d'une matrice symétrique, d'une matrice antisymétrique ? (Voir exercice D.19.)

[12] *Yet more name-dropping.*

4. Montrer que si S est une matrice symétrique de signature (p,q), alors $A \mapsto \mathrm{tr}({}^tASA)$ est de signature (np, nq).

Si $S = \mathrm{I}_n$, on montre que la forme est la norme quadratique $||\cdot||_2$ de l'espace des matrices. Dans le cas général, on peut se ramener au cas où S est la matrice $\mathrm{I}_{(p,q)}$.

Matrices symétriques

D.4. Exercice. *Soit A dans $\mathscr{S}_n^{++}$ et B symétrique.*

1. *Montrer que AB est autoadjointe pour la forme $\langle X, Y\rangle = (A^{-1}X, Y)$,* i.e. *$\langle ABX, Y\rangle = \langle X, ABY\rangle$.*
2. *En déduire que AB est diagonalisable sur $\mathbb{R}$.*

D.5. Exercice. *Soit A, B dans $\mathscr{S}_n^{++}$. Montrer qu'il existe P inversible et D diagonale positive tq $A = {}^tPP$ et $B = {}^tPDP$. En déduire que AB est diagonalisable à valeurs propres positives.*

L'égalité $AB = {}^tPP\,{}^tPDP$ donne $({}^tP)^{-1}AB\,{}^tP = P\,{}^tPDP\,{}^tP$ et donc AB est semblable à une matrice de $\mathscr{S}_n^{++}$.

D.6. Exercice (Un critère pour les matrices symétriques définies positives.). *Soit A une matrice carrée de taille n. Pour tout k de 1 à n, on appelle mineur principal Δ_k de A le mineur correspondant aux k premières lignes et aux k premières colonnes de la matrice A.*

1. *Soit A une matrice (réelle) symétrique définie positive. Montrer que ses mineurs principaux sont tous strictement positifs.*

   Si q est la forme quadratique associée à A dans la base canonique $(e_i)_{1\leqslant i\leqslant n}$, on montrera que Δ_k est le déterminant de la matrice de la restriction de q au sous-espace engendré $(e_i)_{1\leqslant i\leqslant k}$.
2. *Montrer réciproquement que si tous les mineurs principaux sont strictement positifs, alors la matrice A est définie positive.*

   On pourra faire une récurrence sur la dimension de n et utiliser les faits suivants : a) le déterminant d'une matrice est égal au produit de ses valeurs propres, b) une matrice symétrique est définie positive si et seulement si ses valeurs propres sont toutes strictement positives.

D.7. Exercice (Caractérisation des valeurs propres d'une matrice symétrique réelle)

Soit $\mathbb{R}^n$ muni de sa structure euclidienne canonique, notée $(\cdot,\cdot)$. On rappelle qu'une matrice symétrique réelle S est diagonalisable en base orthonormée et que ses valeurs propres sont réelles.

On numérote les valeurs propres λ_i de S de telle sorte que

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n.$$

1. *Soit $||\cdot||$ la norme euclidienne. Montrer que*

$$\lambda_1 = \max_{\|x\|=1} (Sx, x)$$

et que

$$\lambda_n = \min_{\|x\|=1} (Sx, x).$$

Le but de l'exercice est d'obtenir les autres valeurs propres de S de manière similaire.

2. *Soit F_k un sous-espace vectoriel de $\mathbb{R}^n$ de dimension k. Montrer qu'il existe x dans $F_k \cap \mathrm{Vect}(e_k, \ldots, e_n)$ de norme 1. En déduire que*

$$\lambda_k \geqslant \min_{x \in F_k, \|x\|=1} (Sx, x).$$

3. *Montrer que*

$$\lambda_k = \max_{\dim F_k = k} \min_{x \in F_k, \|x\|=1} (Sx, x).$$

Orthogonalité et topologie

D.8. Exercice. *Soit $\mathbb{R}^n$ muni de sa structure euclidienne et k un entier, $0 \leqslant k \leqslant n$. On considère l'application $\perp$: $\mathrm{Gr}_{k,n}(\mathbb{R}) \to \mathrm{Gr}_{n-k,n}(\mathbb{R})$ qui à un sous-espace F de dimension k associe son orthogonal.*

Montrer que cette application est continue pour la topologie des grassmanniennes réelles.

On pourra par exemple fixer un sous-espace F et donc son orthogonal $F^\perp$ et remarquer que l'application de $\mathrm{O}(n)$ dans $\mathrm{Gr}_{n-k,n}(\mathbb{R})$ qui envoie g vers $g(F^\perp)$ passe au quotient sur l'espace homogène $\mathrm{O}(n)/H \simeq \mathrm{Gr}_{k,n}(\mathbb{R})$, où H désigne le stabilisateur de F dans $\mathrm{O}(n)$ (c'est en fait équivalent à dire que si g stabilise $F^\perp$, il stabilise F).

D.9. Exercice. *Soit E un espace vectoriel réel de dimension n et k un entier, $0 \leqslant k \leqslant n$. On considère l'application $\perp$: $\mathrm{Gr}_k(E) \to \mathrm{Gr}_{n-k}(E^*)$ qui à un sous-espace F de dimension k de E associe son orthogonal dans le dual E^*.*

Montrer que cette application est continue pour la topologie des grassmanniennes réelles.

On pourra, soit utiliser l'application linéaire ψ_b du paragraphe A.1.1 et l'exercice précédent, soit faire une preuve directe en recopiant, *mutatis mutandis*, l'exercice précédent...

D.10. Exercice. *Soit q une forme quadratique non dégénérée sur $\mathbb{R}^n$. On considère l'application $\perp : \mathrm{Gr}_{k,n}(\mathbb{R}) \to \mathrm{Gr}_{n-k,n}(\mathbb{R})$ qui à un sous-espace F de dimension k associe son orthogonal $F^\perp$ pour q.*

Montrer que cette application est continue pour la topologie des grassmanniennes réelles.

On pourra utiliser l'exercice précédent et reprendre l'application ψ_b^{-1} du paragraphe A.1.1.

On verra dans le second tome, paragraphe I-2, *que l'application $F \mapsto F^\perp$ est en fait un morphisme de variétés algébriques.*

Autour du théorème de Witt : preuves et conséquences

D.11. Exercice (Gonflement hyperbolique et raffinements). *Soit E un espace vectoriel de dimension n muni d'une forme quadratique non dégénérée q de forme polaire φ, sur un corps de caractéristique différente de 2. On appellera* plan hyperbolique *tout sous-espace P de E de dimension 2 tel que la restriction $q|_P$ a pour matrice $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ dans une base de P.*

Soit F un sous-espace de E. On note F_0 le noyau de $q|_F$, $(x_1, \ldots, x_k)$ une base de F_0 et G un supplémentaire de F_0 dans F.

1. *Soit x un vecteur isotrope de l'espace E et soit G' un sous-espace de dimension $m < n-1$ orthogonal à x. Montrer qu'il existe y orthogonal à G' tel que $\varphi(x,y) \neq 0$. Montrer qu'alors le sous-espace $P := \langle x, y\rangle$ est un plan hyperbolique.*

   On montre par un argument de dimension que l'on peut choisir un y dans $\langle x\rangle^\perp \backslash G'$. On posera ensuite

   $$u = \frac{1}{\varphi(x,y)}\Big(y - \frac{q(y)}{2\varphi(x,y)}x\Big).$$

2. *Montrer que $2k + \dim G \leqslant n$.*

   On a par construction $F = F_0 \oplus G \subset F_0^\perp$, ce qui donne une inégalité des dimensions.

3. *En déduire que l'on peut construire des éléments $y_1, \ldots, y_k$ tels que $P_i := \langle x_i, y_i\rangle$ sont des plans hyperboliques deux à deux orthogonaux et orthogonaux à G.*

   On pourra construire y_i par récurrence comme ci-dessus avec $x = x_i$ et $G' = P_1 \oplus \cdots \oplus P_{i-1} \oplus G$.

4. *Montrer que la restriction de q à $\tilde{F} := P_1 \oplus \cdots \oplus P_k \oplus G$ est non dégénérée.*

   On a une décomposition en sous-espaces orthogonaux tous non dégénérés par construction.

5. *Montrer (sans le théorème de Witt !) que si σ est une isométrie de F vers $\sigma(F) \subset E$, alors σ se prolonge en une isométrie $\tilde{\sigma}$ de $\tilde{F}$ vers $\tilde{\sigma}(\tilde{F})$.*
   Construire un sous-espace non dégénéré contenant $\sigma(F)$.

D.12. Exercice (Théorème de Witt pour des corps classiques). *On veut montrer le théorème de Witt pour les corps $\mathbb{C}$, $\mathbb{R}$, et $\mathbb{F}_q$, Soit E un espace vectoriel de dimension n muni d'une forme quadratique non dégénérée q de forme polaire φ, sur un corps $\mathbb{K}$ de caractéristique différente de 2. Soit F un sous-espace de E de dimension m tel que q soit non dégénérée sur F. Soit F' un sous-espace isométrique à F. On veut montrer que $F^\perp$ et $F'^\perp$ sont isométriques.*

1. *Montrer que $E = F \oplus F^\perp$ et que q est non dégénérée sur $F^\perp$.*
2. *Soit $\mathbb{K} = \mathbb{C}$. Montrer que $F^\perp$ et $F'^\perp$ sont isométriques.*
   Sur $\mathbb{C}$, deux espaces de même dimension munis de formes quadratiques non dégénérées sont isométriques.
3. *Soit $\mathbb{K} = \mathbb{R}$. On suppose que la signature de $q|_F$ est (r_F, s_F) et que la signature de q est (r, s). Montrer que la signature de $q|_{F^\perp}$ est donnée par $(r - r_F, s - s_F)$. En déduire que $F^\perp$ et $F'^\perp$ sont isométriques.*
   Sur $\mathbb{R}$, deux espaces munis de formes quadratiques de même signature sont isométriques.
4. *Soit $\mathbb{K} = \mathbb{F}_q$. On suppose que le discriminant de $q|_F$ est ζ_F et que le discriminant de q est ζ. Montrer que le discriminant de $q|_{F^\perp}$ est $\zeta\zeta_F^{-1}$. En déduire que $F^\perp$ et $F'^\perp$ sont isométriques.*
   Sur $\mathbb{F}_q$, deux espaces de même dimension munis de formes quadratiques non dégénérées de même discriminant sont isométriques.
5. *Pour les corps $\mathbb{C}$, $\mathbb{R}$, $\mathbb{F}_q$, montrer que si σ est une isométrie de F vers F', alors σ se prolonge en une isométrie de E.*
   Soit ι une isométrie entre $F^\perp$ et $F'^\perp$. L'isométrie $\tilde{\sigma} := \sigma \oplus \iota$ de $E = F \oplus F^\perp$ dans $E = F' \oplus F'^\perp$ fait l'affaire.
6. *En déduire le théorème de Witt pour ces trois corps.*
   On se ramène au cas où $q|_F$ est non dégénérée par l'exercice D.11.

D.13. Exercice (Preuve du théorème de Witt. Cas d'une droite non isotrope). *On veut montrer le théorème de Witt 3.4 dans le cadre où F est une droite non isotrope. Soit donc F engendrée par x dans l'espace E muni de la forme quadratique non dégénérée q, de forme polaire associée φ, avec $q(x) \neq 0$. Soit x' tel que $q(x') = q(x)$, engendrant la droite F' et soit σ une isométrie de F vers F' et envoyant x sur x' (elle existe et elle est unique).*

1. *On pose* $z_\epsilon = x + \epsilon x'$, *avec* $\epsilon \in \{1, -1\}$. *Montrer que* $q(z_\epsilon) = 2q(x) + 2\epsilon\varphi(x, x')$. *En déduire qu'il existe une valeur de* ϵ *telle que* $q(z_\epsilon) \neq 0$. *On posera dans la suite* $z = z_\epsilon$ *pour un tel* ϵ *et* $z' = z_{-\epsilon}$.

2. *Soit* τ *l'endomorphisme de* E *tel que*

$$\tau(u) = -\epsilon\big(u - 2\frac{\varphi(u, z)}{q(z)} z\big).$$

Calculer $\tau(u)$ *pour* u *proportionnel à* z *puis pour* u *orthogonal à* z *et en déduire que* τ *est dans* $\mathrm{O}(q)$. *Montrer que* $\tau(z) = \epsilon z$ *et* $\tau(z') = -\epsilon z'$, *puis en sommant que* $\tau(x) = x'$.

On notera que z et z' sont orthogonaux.

3. *Conclure.*

D.14. Exercice (Preuve du théorème de Witt. Cas général)

On veut montrer le théorème de Witt. Dans un premier temps on va montrer par récurrence sur m *que si* F *est un sous-espace non isotrope de dimension* m *de* E, *alors toute isométrie* σ *de* F *sur* $\sigma(F) \subset E$ *se prolonge en une isométrie de* E. *L'initialisation est montrée dans l'exercice* D.13.

1. *On suppose l'hypothèse vraie pour* m. *Montrer que si* F *est non isotrope de dimension* $m + 1$, *alors* F *se décompose en somme orthogonale (et directe) d'une droite* D *et d'un sous-espace non isotrope* F_m *de dimension* m.

   On pourra utiliser le fait qu'il existe une base où la matrice de q_F est diagonale.

2. *Soit* τ_D *l'isométrie de* E *obtenue en appliquant l'exercice* D.13 *à* $\sigma|_D$. *Soit* $\sigma_m := \tau_D^{-1}\sigma$. *Montrer que* σ_m *est une isométrie de* F *sur* $\sigma_m(F)$ *qui fixe les éléments de* D.

3. *Montrer que* σ_m *se restreint en une isométrie de* $F_m \subset D^\perp$ *et que* $\sigma_m(F_m) \subset D^\perp$.

4. *Montrer que* σ_m *se prolonge en une isométrie de* E. *En déduire que* σ *se prolonge en une isométrie de* E.

   Pour la première question, appliquer tout d'abord l'hypothèse de récurrence pour montrer que la restriction de σ_m à F_m se prolonge en une isométrie de $D^\perp$.

5. *Montrer le théorème de Witt dans son cadre général.*

   Si F est isotrope, appliquer le résultat de l'exercice D.11.

D.15. Exercice (SETIM et indice de Witt). *Soit E muni d'une forme quadratique q non dégénérée. On appelle SETIM un sous-espace totalement isotrope,* i.e. *sur lequel q s'annule, et maximal à ce titre.*

1. *Soit F et F' deux SETIM, avec $\dim F \leqslant \dim F'$. Montrer qu'il existe une isométrie de F vers un sous-espace de F'.*
2. *En appliquant le théorème de Witt, montrer que $\nu := \dim F = \dim F'$.*
   Tous les SETIM possèdent donc la même dimension, $\nu = \nu(q)$, appelée indice de Witt de la forme quadratique q.
3. *On appelle* espace hyperbolique *toute somme orthogonale de plans hyperboliques. Montrer que E se décompose en somme orthogonale de H et G, où H est un espace hyperbolique de dimension 2ν et G un sous-espace anisotrope,* i.e. *ne possédant pas de vecteurs isotropes.*
   Utiliser le résultat de l'exercice D.11.
4. *Montrer que si $H \oplus G$ et $H' \oplus G'$ sont deux telles décompositions, alors les espaces G et G' sont isométriques.*
   Appliquer le théorème de Witt à une isométrie que envoie H sur H'.
5. *Quelles sont les dimensions possibles pour G lorsque l'on travaille sur l'un des corps $\mathbb{C}$, $\mathbb{R}$, ou $\mathbb{F}_q$?*
   La dimension ne peut appartenir qu'à $\{0,1\}$ pour $\mathbb{C}$, $\mathbb{N}$ pour $\mathbb{R}$, $\{0,1,2\}$ pour $\mathbb{F}_q$

D.16. Exercice. *Montrer à l'aide du théorème de Witt que si F et F' sont deux sous-espaces isométriques de l'espace vectoriel E muni de la forme quadratique q, alors $F^\perp$ et $F'^\perp$ sont également isométriques.*

Formes bilinéaires alternées

D.17. Exercice (Groupe symplectique et bases symplectiques)
Soit n un entier naturel non nul. On note J la matrice de $\mathscr{M}_{2n}(\mathbb{C})$ définie par

$$J = \begin{pmatrix} 0_n & I_n \\ -I_n & 0_n \end{pmatrix}.$$

On remarque qu'elle satisfait à : $J^{-1} = {}^tJ = J^\star = -J$. On appelle groupe symplectique *le groupe*

$$\mathrm{Sp}_{2n}(\mathbb{C}) = \{P \in \mathscr{M}_{2n}(\mathbb{C}),\ {}^tPJP = J\}.$$

On appelle structure symplectique canonique *de $\mathbb{C}^{2n}$ la forme bilinéaire antisymétrique non dégénérée*

$$\begin{aligned} \sigma : \mathbb{C}^{2n} \times \mathbb{C}^{2n} &\longrightarrow \mathbb{C} \\ (X,Y) &\longmapsto {}^tXJY. \end{aligned}$$

On appelle base symplectique *toute base de la forme*

$$\mathbf{e} = (e_1, e_2, \ldots, e_n, e_{n+1}, e_{n+2}, \ldots, e_{2n})$$

telle que

$$\forall i \in \{1, \ldots, n\},\ \forall j \in \{1, \ldots, 2n\}, \quad \sigma(e_i, e_j) = \delta_{i+n,j}.$$

A. Le groupe $\mathrm{Sp}_{2n}(\mathbb{C})$ et les bases symplectiques

1. *Vérifier que l'on a, pour X et Y quelconques de $\mathbb{C}^n$:*

$$\sigma(X, Y) = {}^tXJY.$$

2. *En déduire qu'une matrice P de $\mathrm{GL}_{2n}(\mathbb{C})$ appartient à $\mathrm{Sp}_{2n}(\mathbb{C})$ si et seulement si elle préserve σ, c'est-à-dire :*

$$\forall X, Y \in \mathbb{C}^{2n}, \quad \sigma(PX, PY) = \sigma(X, Y).$$

3. *Montrer que $\mathrm{Sp}_{2n}(\mathbb{C})$ agit de façon simplement transitive sur les bases symplectiques.*

4. *Montrer que toute famille libre $(e_1, e_2, \ldots, e_p, e_{n+1}, e_{n+2}, \ldots, e_{n+p})$ (avec $1 \leqslant p \leqslant n$), telle que*

$$\forall i \in \{1, \ldots, p\},\ \forall j \in \{1, \ldots, p\} \cup \{n+1, \ldots, n+p\}, \quad \sigma(e_i, e_j) = \delta_{i+n,j},$$

peut être complétée en une base symplectique de $\mathbb{C}^{2n}$. (Une telle famille pourrait être appelée système symplectique.)

B. Groupe $\mathrm{Sp}(n)$ et bases orthonormées symplectiques

On munit également $\mathbb{C}^{2n}$ du produit scalaire hermitien canonique

$$\langle X, Y \rangle = X^*Y.$$

On note $\mathrm{Sp}(n)$ le groupe $\mathrm{Sp}_{2n}(\mathbb{C}) \cap \mathrm{U}(2n)$.

1. *Montrer que pour tout P de $\mathrm{U}(2n)$, on a l'équivalence entre les assertions :*

   (i) $P \in \mathrm{Sp}(n)$;

   (ii) $JP = \overline{P}J$.

2. *On appellera base orthonormée symplectique toute base orthonormée*

$$\mathbf{e} = (e_1, e_2, \ldots, e_n, e_{n+1}, e_{n+2}, \ldots, e_{2n})$$

telle que

$$\forall i \in \{1, \ldots, n\}, \quad e_{n+i} = J\overline{e_i}.$$

Pour $X \in \mathbb{C}^{2n}$, on désigne par S_X le sous-espace vectoriel engendré par les vecteurs X et $J\overline{X}$; on note ${S_X}^{\perp}$ son supplémentaire orthogonal pour le produit scalaire hermitien canonique.

(a) Montrer que si $Y \in S_X{}^{\perp}$, *alors* $J\overline{Y} \in S_X{}^{\perp}$.

(b) En déduire que tout système orthonormé symplectique

$$(e_1, e_2, \ldots, e_p, J\overline{e_1}, J\overline{e_2}, \ldots, J\overline{e_p}), \quad 1 \leqslant p \leqslant n,$$

peut être complété en une base orthonormée symplectique de $\mathbb{C}^{2n}$.

3. *Montrer, en appliquant 1), que pour* $P \in \mathrm{Sp}_{2n}(\mathbb{C})$, *les deux conditions suivantes sont équivalentes :*

(i) $P \in \mathrm{Sp}(n)$;

(ii) les colonnes de P *définissent une base orthonormée symplectique.*

4. *Montrer que* $\mathrm{Sp}(n)$ *agit de façon simplement transitive sur les bases orthonormées symplectiques.*

Instructif

D.18. Exercice (Une inégalité d'Hadamard)

Soit $S = (s_{ij})_{1 \leqslant i,j \leqslant n}$ *une matrice réelle symétrique définie positive. Montrer l'inégalité :* $\det S \leqslant \prod_{i=1}^{n} s_{ii}$.

On pourra construire une forme quadratique q pour laquelle S est une matrice dans une base choisie et considérer le volume du parallélépipède défini par la base.

D.19. Exercice (Quelques formes normales). *Soit* $\big(E, (\cdot\,,\cdot)\big)$ *un espace euclidien de dimension finie.*

1. *Montrer que tout endomorphisme de* E *admet un sous-espace invariant* $W \subset E$ *avec* $\dim W \leqslant 2$.

   Factoriser le polynôme caractéristique.

2. *Soit* f *un endomorphisme de* E *vérifiant*

$$\forall \text{ sous-espace } V \subset E, \quad f(V) \subset V \implies f(V^{\perp}) \subset V^{\perp}.$$

   Montrer qu'il existe une décomposition orthogonale $E = \bigoplus_{1 \leqslant i \leqslant r} W_i$ *en sous-espaces* f*-invariants* W_i, *avec* $\dim W_i \leqslant 2$.

   Procéder par récurrence sur $\dim E$.

3. *En déduire que :*

   (a) tout endomorphisme symétrique est $\mathrm{O}(n)$*-diagonalisable (sur* $\mathbb{R}$*)*[13] ;

[13] On l'a déjà vu dans la proposition 5.3.

(b) la matrice de tout endomorphisme orthogonal est $\mathrm{O}(n)$*-semblable à une matrice de la forme*

$$\begin{pmatrix} \mathbf{1}_{n_+} & 0 & 0 & \cdots & 0 \\ 0 & -\mathbf{1}_{n_-} & 0 & \cdots & \vdots \\ 0 & 0 & R_1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & R_r \end{pmatrix}, \quad R_i = \begin{pmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{pmatrix}, \ \theta_i \in \mathbb{R};$$

(c) la matrice de tout endomorphisme antisymétrique est $\mathrm{O}(n)$*-semblable à une matrice de la forme*

$$\begin{pmatrix} \mathbf{0}_{n_0} & 0 & \cdots & 0 \\ 0 & A_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_r \end{pmatrix}, \quad A_i = \begin{pmatrix} 0 & \theta_i \\ -\theta_i & 0 \end{pmatrix}, \ \theta_i \in \mathbb{R}.$$

Utiliser la proposition 5.4.

D.20. Exercice (Trigonalisation en base orthonormée et continuité du rayon spectral)

Soit $\mathbb{C}^n$ *muni de sa structure hermitienne canonique. On rappelle que le rayon spectral* $\rho(A)$ *d'une matrice* A *de* $\mathscr{M}_n(\mathbb{C})$ *est égal au maximum des modules de ses valeurs propres.*

1. *On note* φ *l'endomorphisme de* $\mathbb{C}^n$ *associé à* A *dans la base canonique. Montrer que* $\mathbf{e} = (e_1, \ldots, e_n)$ *est une base pour laquelle* φ *est triangulaire si et seulement si tous les sous-espaces* $F_i = \langle e_1, \ldots, e_i \rangle$ *sont stables par* φ*.*

2. *En appliquant le procédé de Gram-Schmidt, déduire que* φ *est trigonalisable dans base orthonormée.*

3. *Montrer que pour l'action de* $\mathrm{U}_n(\mathbb{C})$ *par conjugaison sur* $\mathscr{M}_n(\mathbb{C})$*, toute orbite contient une matrice triangulaire.*

4. *Montrer que l'application* $\rho : \mathscr{M}_n(\mathbb{C}) \to \mathbb{R}^+$ *est continue.*

   On pourra tout d'abord montrer la continuité sur les matrices triangulaires, puis montrer que si A_k converge vers A, alors $\rho(A_k)$ est bornée, puis en utilisant la question 3, que toute valeur d'adhérence de cette suite est égale à $\rho(A)$.

D.21. Exercice (Formes de Hankel)

Le but de l'exercice est de construire une forme quadratique réelle s associée à un polynôme réel P et telle que la signature de s permette de calculer le nombre de racines distinctes ainsi que le nombre de racine réelles distinctes de P.

Soit donc P un polynôme réel de degré n et soit $x_1, \dots, x_t$ ses racines. Supposons que x_k est de multiplicité m_k. On note $s_k = m_1 x_1^k + \cdots + m_t x_t^k$ les sommes de Newton (et l'on pose $s_0 = n$).

1. *Montrer que $\sum_{0 \leqslant i,j \leqslant n-1} s_{i+j} X_i X_j$ définit une forme quadratique s sur $\mathbb{C}^n$ et qu'elle définit une forme quadratique $s_{\mathbb{R}}$ sur $\mathbb{R}^n$.*
2. *Soit φ_k la forme linéaire sur $\mathbb{C}^n$ donnée par*

$$\varphi_k(X_0, \dots, X_{n-1}) = X_0 + x_k X_1 + x_k^2 X_2 + \cdots + x_k^{n-1} X_{n-1}$$

   et soit (p, q) la signature de $s_{\mathbb{R}}$.

   (a) Montrer par un déterminant de Vandermonde que les φ_k sont indépendants.

   (b) Montrer que $s = \sum_{k=1}^{t} m_k \varphi_k^2$.

   (c) En déduire que le nombre t de racines distinctes de P est $p + q$.
3. *Montrer que la signature de $\varphi_k^2 + \bar{\varphi}_k^2$ vaut $(1, 0)$ si x_k est réel et $(1, 1)$ sinon.*
4. *En déduire que le nombre de racines réelles distinctes de P est égal à $p - q$.*

D.22. Exercice (Espace des SETIM)

On considère un espace vectoriel réel E de dimension $2n$, $n \geqslant 1$, muni d'une forme quadratique q de signature (n, n).

A. *Le but de cette partie est d'étudier l'action du groupe orthogonal sur l'ensemble des sous-espaces totalement isotropes maximaux pour l'inclusion.*

1. *Montrer l'assertion suivante :*

$$F \text{ est un espace totalement isotrope} \iff F \subset F^{\perp}.$$

   En déduire que si F est totalement isotrope, alors $\dim F \leqslant n$.
2. *Soit F, F' deux sous-espaces totalement isotropes de E. Montrer l'assertion suivante :*

$$\exists\, g \in \mathrm{O}(q),\ g(F) = F' \iff \dim F = \dim F'.$$

   Penser à utiliser le théorème de Witt.

3. (Rappel) *On appelle SETIM un sous-espace totalement isotrope maximal et l'on note $\mathscr{I}$ l'ensemble de tous les SETIM de E.*

   (a) Montrer qu'il existe une base $\mathscr{B}$ dans laquelle q admet pour matrice

$$\mathrm{mat}_{\mathscr{B}}(q) = \begin{pmatrix} 0_n & \mathrm{I}_n \\ \mathrm{I}_n & 0_n \end{pmatrix},$$

   où 0_n et I_n désignent respectivement la matrice nulle et la matrice identité $n\times n$. On fixera $\mathscr{B}$ dans la suite.

   (b) Soit Λ le sous-espace engendré par les n premiers vecteurs de $\mathscr{B}$. Montrer que $\Lambda \in \mathscr{I}$.

   (c) Montrer alors que tout SETIM de E est de dimension n.

   C'est encore une application du théorème de Witt.

4. *Montrer que le groupe $\mathrm{O}(q)$ agit sur $\mathscr{I}$ par $g \cdot L := g(L)$ et que cette action est transitive.*

B. *Dans cette partie, on veut étudier les propriétés topologiques de $\mathscr{I}$ dans la n-grassmannienne de E.*

1. *Soit g dans $\mathrm{GL}(E)$ tel que*

$$\mathrm{mat}_{\mathscr{B}}(g) = \begin{pmatrix} A & C \\ B & D \end{pmatrix},$$

   où A, B, C, D sont des blocs de taille $n\times n$. Montrer que le groupe d'isotropie $\mathrm{O}(q)_\Lambda$ est défini par la condition :

$$g \in \mathrm{O}(q)_\Lambda \;\Leftrightarrow\; A \in \mathrm{GL}_n(\mathbb{C}),\ B = 0,\ C = H({}^tA^{-1}),\ D = {}^tA^{-1},$$

   où H est une matrice antisymétrique complexe de taille $n\times n$ quelconque.

2. *En déduire que le groupe $\mathrm{O}(q)_\Lambda$ est isomorphe à un produit semi-direct du groupe $\mathrm{GL}_n(\mathbb{R})$ par le groupe additif $\mathbb{R}^{n(n-1)/2}$, pour une action que l'on précisera.*

3. *Montrer que $\mathscr{I}$ est un compact à deux composantes connexes pour la topologie induite de la grassmannienne.*

   On pourra considérer $\mathscr{I}$ comme un espace homogène. Pour la première assertion, on pourra utiliser le fait que $\mathrm{O}(q)$ est un fermé dans $\mathrm{GL}_{2n}(\mathbb{R})$, puis que la surjection canonique de $\mathrm{GL}_{2n}(\mathbb{R})$ sur le quotient $\mathrm{GL}_{2n}(\mathbb{R})/\mathrm{GL}_{2n}(\mathbb{R})_\Lambda$ envoie un fermé sur un fermé. Pour la seconde assertion, on montrera que la fonction déterminant passe au quotient du stabilisateur dans $\mathrm{O}(q)$. On rappelle aussi que le groupe spécial orthogonal sur $\mathbb{R}$ est connexe.

4. *Reconnaître, dans le cas où $n = 2$, les deux composantes connexes de l'espace des SETIM, sur la figure ci-dessous, représentant la quadrique d'équation $x^2 + y^2 - z^2 - t^2 = 0$ dans l'espace projectif $\mathbb{P}^3(\mathbb{R})$.*

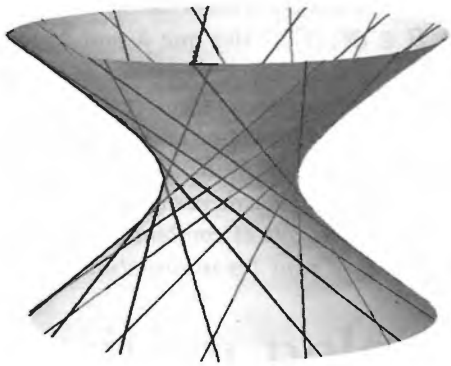


Figure D.1. *SETIM dans la quadrique* $x^2 + y^2 - z^2 - t^2 = 0$

D.23. Exercice (Inégalités de Weyl)

Pour toute matrice symétrique réelle S *de taille* n*, on note* $\lambda_k(S)$ *sa* k*-ième valeur propre, où toutes les valeurs propres de* S *sont ordonnées par ordre décroissant.*

Soit S *et* S' *deux matrices symétriques réelles de taille* n*.*

1. *Question préliminaire. Soit* F*,* G *et* H *des sous-espaces vectoriels de* $\mathbb{R}^n$ *tels que* $\dim(F) + \dim(G) + \dim(H) > 2n$*. Montrer que* $F \cap G \cap H$ *n'est pas réduit à* $\{0\}$*.*

   Considérer l'application linéaire de $F \times G \times H$ dans $E \times E$ définie par $(f, g, h) \mapsto (f - h, g - h)$.

2. *Montrer que l'on a :*
$$\lambda_1(S + S') \leqslant \lambda_1(S) + \lambda_1(S').$$

3. *Soit* i *et* j *deux entiers tels que* $1 \leqslant i, j, i + j - 1 \leqslant n$*. Montrer que l'on a :*
$$\lambda_{i+j-1}(S + S') \leqslant \lambda_i(S) + \lambda_j(S').$$

   On fixera une base de vecteurs propres $\big(v_k(S)\big)$, resp. $\big(v_k(S')\big)$, associée aux valeurs propres $\lambda_k(S)$, resp. $\lambda_k(S')$. Pour tout sous-espace de dimension $i + j - 1$, on considérera son intersection, non triviale, avec $\langle v_i(S), \ldots, v_n(S)\rangle \cap \langle v_j(S'), \ldots, v_n(S')\rangle$ et l'on appliquera le résultat de l'exercice D.7.

L'exercice se généralise sans peine en remplaçant les matrices symétriques par des matrices hermitiennes.

Les inégalités de Weyl sont des conditions nécessaires sur un triplet (λ, μ, ν) pour qu'il existe $A, B \in \mathscr{H}_n(\mathbb{C})$ telles que λ soit le spectre de A, μ celui de B et ν celui de $A + B$. En 1962, Horn les a complétées par d'autres inégalités du même type et a conjecturé que le système obtenu donnait des conditions suffisantes. C'est le problème de Horn, *résolu en 1998 par A. A. Klyachko d'une part, en 1999 par A. Knutson et T. Tao d'autre part. On trouvera une présentation accessible de ces questions mêlant géométrie algébrique, théorie des représentations et combinatoire, dans l'article [7]. Nous remercions Nicolas Ressayre pour les informations sur le sujet.*

Chapitre VI

Décomposition polaire et applications

Toute personne qui s'est intéressée un temps à la politique ou ayant pratiqué un quelconque jeu de rôle aura compris l'adage « diviser pour régner ». Cette tactique étant féconde dans des milieux divers allant des sciences humaines à l'informatique, il est naturel qu'elle le soit en mathématiques. Les théorèmes de décomposition en sont l'illustration flagrante. Par exemple, la décomposition de Dunford permet de séparer les problèmes de réduction en traitant d'un côté le cas diagonalisable (avantages : orbites fermées, polynôme caractéristique comme invariant total...) et de l'autre le cas nilpotent (avantages : nombre fini d'orbites, une seule valeur propre...).

Nous allons voir que, pour la plupart des groupes classiques sur $\mathbb{R}$ et $\mathbb{C}$, ceux-ci se divisent en deux parties aux avantages complémentaires. Une partie est compacte, l'autre ne l'est pas, mais en revanche est homéomorphe à un espace topologique particulièrement sympathique : un espace vectoriel de dimension finie. C'est donc la décomposition polaire qui va permettre de régner sur les groupes classiques, au moins de façon topologique. Comme application, nous donnerons une description des composantes connexes des groupes $\mathrm{O}(p,q)$, puis des quadriques réelles sur lesquelles les $\mathrm{O}(p,q)$ agissent transitivement.

1. Théorème de décomposition polaire

On commence par des rappels et des notations.

1.1. Rappel. Voir la proposition V-1.6.

Soit n un entier naturel. L'ensemble des matrices *symétriques définies positives* de taille $n \times n$ est :

$$\begin{aligned}\mathscr{S}_n^{++}(\mathbb{R}) &= \{S \in \mathrm{GL}_n(\mathbb{R}) \,:\, {}^tS = S;\ \forall x \in \mathbb{R}^n \setminus \{0\},\ {}^txSx > 0\} \\ &= \{P\,{}^tP \in \mathscr{M}_n(\mathbb{R}) \,:\, P \in \mathrm{GL}_n(\mathbb{R})\}.\end{aligned}$$

Pour $x \in \mathbb{C}^n$, on note $x^* = {}^t\overline{x}$ le vecteur adjoint de x ; pour $H \in \mathscr{M}_n(\mathbb{C})$, on note $M^* = {}^t\overline{H}$ la matrice adjointe de H. L'ensemble des matrices *hermitiennes définies positives* $n \times n$ est :

$$\begin{aligned}\mathscr{H}_n^{++}(\mathbb{C}) &= \{H \in \mathrm{GL}_n(\mathbb{C}) \,:\, H^* = H;\ \forall x \in \mathbb{C}^n \setminus \{0\},\ x^*Hx > 0\} \\ &= \{PP^* \in \mathscr{M}_n(\mathbb{C}) \,:\, P \in \mathrm{GL}_n(\mathbb{C})\}.\end{aligned}$$

Avant d'attaquer le théorème de décomposition polaire, on pourra remarquer que l'ensemble des matrices symétriques définies positives forment un espace homogène.

1.2. Exercice. *Montrer que l'action de congruence (réelle donc) du groupe linéaire* $\mathrm{GL}_n(\mathbb{R})$ *sur* $\mathscr{S}_n^{++}$ *permet de réaliser* $\mathscr{S}_n^{++}$ *comme quotient de* $\mathrm{GL}_n(\mathbb{R})$ *par le groupe orthogonal* $\mathrm{O}(n, \mathbb{R})$.

Voici ce qui peut donc être vu comme un relèvement de cette propriété (et son analogue sur $\mathbb{C}$) et qui sera le théorème principal du chapitre.

1.3. Théorème (Décomposition polaire). *La multiplication matricielle induit des homéomorphismes :*

(i) $\mu : \mathrm{O}(n, \mathbb{R}) \times \mathscr{S}_n^{++}(\mathbb{R}) \xrightarrow{\cong} \mathrm{GL}_n(\mathbb{R}),\ (O, S) \longmapsto OS$;

(ii) $\mu : \mathrm{U}(n, \mathbb{C}) \times \mathscr{H}_n^{++}(\mathbb{C}) \xrightarrow{\cong} \mathrm{GL}_n(\mathbb{C}),\ (U, H) \longmapsto UH$.

Démonstration. (i) L'application μ est bien définie et continue. Montrons sa surjectivité. Soit $M \in \mathrm{GL}_n(\mathbb{R})$. La matrice ${}^tMM$ appartient à $\mathscr{S}_n^{++}(\mathbb{R})$, car elle est dans l'orbite de I_n pour l'action de congruence (proposition V-1.6). On peut donc diagonaliser ${}^tMM$ dans une base orthonormée :

$${}^tMM = P \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} P^{-1},$$

pour $P \in \mathrm{O}(n,\mathbb{R})$ convenable, et $\lambda_i > 0$ pour tout i. On pose alors :

$$S = P \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix} P^{-1} :$$

c'est une matrice symétrique, puisque P est orthogonale, et définie positive, car ses valeurs propres sont strictement positives.

On a : $S^2 = {}^tMM$ et, si l'on pose : $O = MS^{-1}$, il vient :

$${}^tOO = {}^t(MS^{-1})MS^{-1} = {}^tS^{-1}\,{}^tMMS^{-1} = S^{-1}S^2S^{-1} = \mathrm{I}_n.$$

Ainsi : $M = OS$, où $O \in \mathrm{O}(n,\mathbb{R})$ et $S \in \mathscr{S}_n^{++}(\mathbb{R})$, donc μ est surjective.

Montrons que μ est injective. Supposons que l'on ait : $M = OS = O'S'$, avec O' dans $\mathrm{O}(n,\mathbb{R})$ et S' dans $\mathscr{S}_n^{++}(\mathbb{R})$.

Alors, il vient : $S^2 = {}^tMM = {}^t(O'S')O'S' = {}^tS'\,{}^tO'O'S'$ d'où : $S^2 = S'^2$.

Soit Q un polynôme[1] tel que pour tout i, $Q(\lambda_i) = \sqrt{\lambda_i}$. Alors :

$$S = P \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix} P^{-1} = PQ\left(\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}\right) P^{-1}$$

$$= Q\left(P \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} P^{-1}\right) = Q(S^2) = Q(S'^2).$$

Or, S' commute avec S'^2, donc avec $Q(S'^2) = S$, et donc S' et S sont diagonalisables dans une base commune. Il existe ainsi une matrice de passage P_0 qui permet de les diagonaliser simultanément.

Ainsi, si l'on note

$$\mathrm{diag}(\mu_i) = \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix},$$

[1] On peut utiliser un polynôme d'interpolation de Lagrange : quitte à rénuméroter les λ_i, on peut supposer que $\lambda_1, \ldots, \lambda_r$ sont distincts et que $\lambda_{r+1}, \ldots, \lambda_n$ apparaissent parmi ceux-là. On prend alors :

$$Q(X) = \sum_{i=1}^{r} \sqrt{\lambda_i} \prod_{\substack{1 \leqslant j \leqslant r \\ j \neq i}} \frac{X - \lambda_j}{\lambda_i - \lambda_j}.$$

on a : $S' = P_0 \operatorname{diag}(\mu'_i) P_0^{-1}$ et $S = P_0 \operatorname{diag}(\mu_i) P_0^{-1}$. Par suite :

$$\begin{aligned} S'^2 = S^2 &\Longrightarrow P_0 \operatorname{diag}(\mu_i'^2) P_0^{-1} = P_0 \operatorname{diag}(\mu_i^2) P_0^{-1} \\ &\Longrightarrow \forall 1 \leqslant i \leqslant n,\ \mu_i'^2 = \mu_i^2 \\ &\Longrightarrow \forall 1 \leqslant i \leqslant n,\ \mu'_i = \mu_i \quad (S, S' \in \mathscr{S}_n^{++} \text{ donne } \mu_i, \mu'_i \in \mathbb{R}^{+*}) \\ &\Longrightarrow S = S'. \end{aligned}$$

Ainsi, on a : $O = O'$, d'où finalement, l'injectivité de μ.

Il reste à montrer que μ^{-1} est continue. Soit $(M_p)_{p\in\mathbb{N}}$ une suite de $\mathrm{GL}_n(\mathbb{R})$ qui converge vers M. On note, pour tout entier p : $(O_p, S_p) = \mu^{-1}(M_p)$, de sorte que $M_p = O_p S_p$, et $\mu^{-1}(M) = (O, S) \in \mathrm{O}(n, \mathbb{R}) \times \mathscr{S}_n^{++}(\mathbb{R})$. On montre que les suites (O_p) et (S_p) convergent respectivement vers O et S.

On sait par la proposition II-4.1.6 que $\mathrm{O}(n, \mathbb{R})$ est compact. Soit $\overline{O}$ une valeur d'adhérence de $(O_p)_{p\in\mathbb{N}}$, *i.e.* telle que O_{p_k} converge vers $\overline{O}$ pour une sous-suite $(O_{p_k})_{k\in\mathbb{N}}$ de (O_p). Alors, la sous-suite S_{p_k} converge vers $\overline{O}^{-1} M$, matrice symétrique et définie positive, car :

$$\overline{S} := \overline{O}^{-1} M \in \mathrm{GL}_n(\mathbb{R}) \cap \overline{\mathscr{S}_n^{++}(\mathbb{R})} = \mathrm{GL}_n(\mathbb{R}) \cap \mathscr{S}_n^{+}(\mathbb{R}) = \mathscr{S}_n^{++}(\mathbb{R}).$$

(Ah ! Les adhérences d'orbites...) On a donc, par unicité de la décomposition polaire (injectivité de μ) :

$$M = \overline{O}\,\overline{S} \quad \text{et} \quad \overline{O} = O \quad \text{et} \quad \overline{S} = S.$$

(ii) Similaire. Laissé en exercice. □

1.4. Exemple. Notons que pour $n = 1$, les matrices hermitiennes définies positives sont les réels strictement positifs ($\mathscr{H}_1^{++} = \mathbb{R}^{+*}$) et les unitaires sont les complexes de module 1. En particulier, l'assertion (ii) exprime la décomposition $z = \rho e^{i\theta}$, unique pour tout z non nul. Ceci explique certainement l'appellation contrôlée « décomposition polaire ».

1.5. Remarque. De la preuve du théorème, on retient un résultat intéressant : pour tout matrice symétrique réelle définie positive A, il existe une unique matrice réelle symétrique définie positive B telle que $B^2 = A$. Dit autrement, $\mathscr{S}_n^{++}(\mathbb{R})$ est munie d'une application « racine carrée » que l'on notera $S \mapsto \sqrt{S}$. Il en est de même dans $\mathscr{H}_n^{++}(\mathbb{R})$..

1.6. Corollaire. *Pour toute matrice inversible de $\mathscr{M}_n(\mathbb{R})$, on a :*

$$\|A\|_2 = \sqrt{\rho({}^tAA)},$$

où ρ est le rayon spectral donné par :

$$\rho(M) = \max_{1 \leqslant i \leqslant n} \{|\lambda_i| \,:\, \lambda_i \in \mathrm{Spec}(M)\}.$$

Démonstration. Soit $A = OS$. Comme $||OS(x)||_2 = ||S(x)||_2$ pour tout vecteur x, on a : $\|A\|_2 = \|S\|_2$. Comme S est symétrique réelle, elle est diagonalisable dans une base orthonormée $(v_1, \ldots, v_n)$, ordonnée de sorte que les valeurs propres correspondantes soient : $|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_n|$. Maintenant, si $x := \sum_i x_i v_i$ est de norme 1, on a : $||S(x)||_2 = ||\sum_i \lambda_i x_i v_i||_2 \leqslant \rho(S) = |\lambda_1|$, la borne étant atteinte pour $x = v_1$.

On a montré l'égalité : $\|S\|_2 = \rho(S)$ et l'on peut conclure par :

$$\rho(S) = |\lambda_1| = \sqrt{|\lambda_1^2|} = \sqrt{\rho(S^2)} = \sqrt{\rho({}^t AA)}.$$

□

1.7. Remarque. Grâce à la continuité du rayon spectral (étudiée dans l'exercice V-D.20) et à la densité des matrices inversibles complexes, l'égalité du corollaire est encore valable pour toute matrice A.

La décomposition polaire a aussi comme corollaire la maximalité du groupe orthogonal $\mathrm{O}(n) = O_n(\mathbb{R})$ parmi les sous-groupes compacts de $\mathrm{GL}_n(\mathbb{R})$.

1.8. Corollaire (Maximalité du groupe orthogonal). *Tout sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$ qui contient le groupe orthogonal $\mathrm{O}(n)$ est le groupe $\mathrm{O}(n)$ lui-même.*

Démonstration. Soit G un groupe compact contenant $\mathrm{O}(n)$ et soit g dans G, on peut écrire : $g = os$, avec o dans $\mathrm{O}(n)$ et $s \in \mathscr{S}_n^{++}$. Les hypothèses impliquent que $s = o^{-1}g \in G$ et donc $s^k \in G$ pour tout k dans $\mathbb{Z}$. Or, s est diagonalisable et son spectre est contenu dans $\mathbb{R}^{+*}$. Comme G est compact, il vient en utilisant le corollaire ci-dessus que les valeurs propres de s ne peuvent être strictement supérieures à 1 (sinon la suite $(\|s^k\|_2)_{k\in\mathbb{N}}$ ne serait pas bornée) ni strictement inférieures à 1 (sinon la suite $(\|s^{-k}\|_2)_{k\in\mathbb{N}}$ ne serait pas bornée). Donc, $s = \mathrm{Id}$ et $g = o \in \mathrm{O}(n)$. □

1.9. Remarque. Les groupes classiques G réels ou complexes que nous étudions contiennent un sous-groupe compact K maximal et tel que G/K soit homéomorphe à un espace vectoriel de dimension finie.

Ce sous-groupe est unique à conjugaison près : c'est $\mathrm{O}(n)$ dans $\mathrm{GL}_n(\mathbb{R})$, $\mathrm{SO}(n)$ dans $\mathrm{SL}_n(\mathbb{R})$ par exemple, ou encore $\mathrm{U}(n)$ dans $\mathrm{GL}_n(\mathbb{C})$ et $\mathrm{SU}(n)$ dans $\mathrm{SL}_n(\mathbb{C})$. (Mais attention, $\mathrm{O}(n, \mathbb{C})$ n'est pas compact : par exemple, il agit transitivement sur la nappe $z_1^2 + z_2^2 = 1$ qui contient l'hyperbole réelle non bornée $x_1^2 - y_2^2 = 1$.)

Cette remarque est liée à un résultat important sur lequel il serait bien de s'attarder. Tout sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$ est conjugué à un sous-groupe du groupe orthogonal $\mathrm{O}(n)$. Nous ne le démontrerons pas, mais nous renvoyons le lecteur au paragraphe 3.6.5 de [25]. En revanche, on comprend très bien l'idée de la preuve en regardant le *toy model* des groupes compacts, c'est-à-dire les groupes finis.

1.10. Proposition. *Tout sous-groupe fini G de $\mathrm{GL}_n(\mathbb{R})$ est conjugué à un sous-groupe de $\mathrm{O}(n)$.*

Démonstration. Comme on a vu que sur $\mathbb{R}$ toutes les formes quadratiques définies positives sont dans la même orbite de congruence et que tous les stabilisateurs d'une même orbite sont conjugués, il suffit de montrer que G est dans le stabilisateur d'une forme quadratique définie positive q.

Pour cela, considérons la forme euclidienne canonique $\langle \cdot\,, \cdot \rangle$ de $\mathbb{R}^n$ et posons

$$q(x) = \sum_{g \in G} \langle g(x), g(x) \rangle.$$

C'est bien d'une forme quadratique définie positive. De plus, si $h \in G$, on a :

$$q\big(h(x)\big) = \sum_{g \in G} \langle g\big(h(x)\big), g\big(h(x)\big) \rangle = \sum_{k \in G} \langle k(x), k(x) \rangle,$$

à l'aide du changement de variable $k = gh$. Cela prouve que $h \in \mathrm{O}(q)$ comme voulu, puisque $\mathrm{O}(q)$ est conjuguée à $\mathrm{O}(n)$. □

1.11. Remarque. On en profite pour observer cette idée féconde qui consiste à introduire une forme G-invariante lorsque G est un groupe fini (on peut le généraliser dans le cas compact en remplaçant la somme par une certaine intégrale). Attention tout de même : sur un corps quelconque, il n'est pas clair que cette forme ne soit pas nulle. C'est le fait d'être sur les réels qui assure, par positivité, qu'elle reste non dégénérée. Sur le corps des complexes, on adapte la construction avec une forme hermitienne.

1.12. Remarque. On peut imaginer à l'aide de cette preuve comment étendre ce résultat (et beaucoup d'autres) d'un groupe fini à un groupe compact : il suffit de remplacer la somme finie sur G par une intégrale sur G, à condition d'avoir sur G une mesure G-invariante qui remplace la mesure de comptage des groupes finis. L'existence d'une telle mesure invariante est un fait général dans les groupes topologiques localement compacts et s'établit ici plus simplement en tenant compte de la compacité de G et d'un théorème de point fixe.

2. L'exponentielle

L'exponentielle est un outil fondamental dans le cadre des groupes de Lie, *cf.* chapitre IX, p. 267 : dans ce cadre, elle permet (presque) de reconstruire le groupe à partir de son espace tangent en l'identité. Le premier exemple basique bien connu est la formule de Moivre $\exp(it) = \cos t + i \sin t$ $(t \in \mathbb{R})$, où l'on voit que le groupe du cercle peut être retrouvé à partir de son espace tangent en 1 (l'axe des imaginaires), et ce passage se fait via l'exponentielle comme on peut le voir sur la figure 2.1.

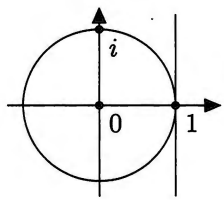


Figure 2.1. *Groupe du cercle unité et son algèbre de Lie*

Soit $\mathbb{K}$ le corps des réels ou des complexes. On définit :

$$\begin{aligned} \exp : \mathscr{M}_n(\mathbb{K}) &\longrightarrow \mathscr{M}_n(\mathbb{K}) \\ A &\longmapsto \exp(A) = \sum_{i=0}^{\infty} \frac{1}{i!} A^i. \end{aligned}$$

On comprendra sans mal que si l'on pose $n = 1$, on retrouve la fonction exponentielle bien connue sur $\mathbb{R}$ et sur $\mathbb{C}$.

Voici un résumé des diverses formules et propriétés de l'exponentielle qui serviront dans la suite. Nous laissons les preuves au lecteur en donnant toutefois quelques indications.

2.1. Proposition.

(i) *La série qui définit l'exponentielle converge normalement sur tout compact de* $\mathscr{M}_n(\mathbb{C})$. *L'exponentielle est une fonction continue.*

(ii) *Pour* A *quelconque et* P *inversible, on a :* $\exp(PAP^{-1}) = P \exp(A) P^{-1}$ *et* $\exp({}^tA) = {}^t\exp(A)$.

(iii) *Si* A *et* B *commutent, on a :* $\exp(A+B) = \exp(A)\exp(B)$.

(iv) *Pour toute matrice* A, *on a :* $\exp(-A)\exp(A) = \mathrm{Id} = \exp(A)\exp(-A)$. *En particulier,* $\exp(A)$ *est inversible.*

(v) *Pour une matrice diagonale, on a :*

$$\exp \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix} = \begin{pmatrix} e^{a_1} & & 0 \\ & \ddots & \\ 0 & & e^{a_n} \end{pmatrix} \quad (a_1, \ldots, a_n \in \mathbb{C}).$$

(vi) *Si une matrice N est nilpotente, alors* $\exp(N) - \mathrm{Id}$ *est nilpotente.*

(vii) *Le spectre de* $\exp(A)$ *est* $\{e^{\lambda},\ \lambda \in \mathrm{Spec}(A)\}$.

(viii) *Pour A quelconque,* $\{\exp(tA),\ t \in \mathbb{R}\}$ *est un sous-groupe de* $\mathrm{GL}_n(\mathbb{K})$.

(ix) *La fonction* exp *n'est pas injective.*

(x) *L'exponentielle est différentiable en la matrice nulle* 0 *et sa différentielle est l'identité* Id *de* $\mathscr{M}_n(\mathbb{C})$ *; c'est un difféomorphisme local sur un voisinage de* 0.
Plus précisément, on peut définir son inverse sur la boule ouverte centrée en I_n *de rayon* 1 *de n'importe quelle norme subordonnée* $||\cdot||$ *: si* $||H|| < 1$, *on pose :*

$$\exp^{-1}(\mathrm{I}_n + H) = \log(\mathrm{I}_n + H) = \sum_{k \geqslant 1} \frac{(-1)^{k-1}}{k} H^k.$$

Esquisse. Pour l'assertion (iii), le membre de droite est un produit de séries $\sum_{i=1}^{\infty} a_i \sum_{j=1}^{\infty} b_j$ et le membre de gauche est, puisque A et B commutent, la série-produit $\sum_{k=1}^{\infty} \sum_{i=1}^{\infty} a_i b_{k-i}$ de ces deux séries. La convergence absolue assure que la série-produit converge vers le produit des séries. Pour l'assertion (vii), on peut utiliser par exemple la décomposition de Dunford. Pour l'assertion (ix), on peut calculer l'exponentielle de

$$\begin{pmatrix} 0 & 2k\pi \\ -2k\pi & 0 \end{pmatrix},\ k \in \mathbb{Z}.$$

Dans l'assertion (x), l'existence de l'inverse résulte du théorème d'inversion locale IX-A.1. Pour sa décomposition en série, il s'agit d'une utilisation classique *via* les séries formelles : on vérifie que les séries formelles de $\mathbb{C}[[X]]$ qui définissent $\exp(X) - 1$ et $\log(1 + X)$ sont bien inverses l'une de l'autre pour la composition des séries formelles. (Attention à ne pas oublier de montrer qu'elles sont composables.) □

2.2. Remarque. L'exponentielle permet de transformer la décomposition de Dunford additive $A = D + N$ de A dans $\mathscr{M}_n(\mathbb{K})$, en la décomposition de Dunford multiplicative de $\exp(A)$ dans $\mathrm{GL}_n(\mathbb{K})$, $exp(A) = \exp(D)\exp(N)$ (remarque III-A.7). Aneffet : si D est diagonalisable, $\exp(D)$ est aussi diagonalisable et si N est nilpotente, $\exp(N)$ est unipotente ($\mathrm{I}_n - \exp N$ est nilpotente). De plus, D et N commutent, donc leurs exponentielles aussi.

2.3. Proposition

(i) *L'application* $\exp : \mathscr{S}_n(\mathbb{R}) \longrightarrow \mathscr{S}_n^{++}(\mathbb{R})$ *est un homéomorphisme.*

(ii) *L'application* $\exp : \mathscr{H}_n(\mathbb{C}) \longrightarrow \mathscr{H}_n^{++}(\mathbb{C})$ *est un homéomorphisme.*

Démonstration. (i) Soit $S \in \mathscr{S}_n(\mathbb{R})$. On peut écrire : $S = P\,\mathrm{diag}(\lambda_i)P^{-1}$, où P est orthogonale et $\mathrm{diag}(\lambda_i)$ est une matrice diagonale dont les coefficients sont les $\lambda_i \in \mathbb{R}$. On a :

$$\exp(S) = P\,\mathrm{diag}(e^{\lambda_i})P^{-1} = P\,\mathrm{diag}(e^{\lambda_i})\,{}^tP \in \mathscr{S}_n^{++}(\mathbb{R}),$$

car $e^{\lambda_i} > 0$ pour tout i et P est orthogonale, donc l'exponentielle envoie bien $\mathscr{S}_n(\mathbb{R})$ dans $\mathscr{S}_n^{++}(\mathbb{R})$ et, par restriction, elle reste continue.

Montrons la surjectivité. Soit $B \in \mathscr{S}_n^{++}(\mathbb{R})$. On a donc $B = P\,\mathrm{diag}(\mu_i)P^{-1}$, avec $P \in \mathrm{O}(n,\mathbb{R})$. Alors, puisque $\mu_i > 0$, la matrice $A = P\,\mathrm{diag}(\ln \mu_i)P^{-1}$ est symétrique et l'on a : $\exp A = B$, d'où la surjectivité.

Montrons l'injectivité. Supposons que $\exp A = \exp A'$, avec $A, A' \in \mathscr{S}_n(\mathbb{R})$. Soit Q un polynôme interpolateur de Lagrange tel que $Q(e^{\lambda_i}) = \lambda_i$ pour tout i. Alors, A' commute avec $Q(\exp A') = Q(\exp A) = A$. Par diagonalisation simultanée, on conclut, comme dans la preuve de la décomposition polaire, que $A = A'$ et donc exp est injective.

Montrons la bicontinuité. Soit $(B_p)_{p\in\mathbb{N}} = \big(\exp(A_p)\big)_{p\in\mathbb{N}}$ une suite de $\mathscr{S}_n^{++}(\mathbb{R})$ qui converge vers $B = \exp A \in \mathscr{S}_n^{++}(\mathbb{K})$. Montrons que (A_p) converge vers A.

Comme elle converge, la suite (B_p) est bornée pour $\|\cdot\|_2$. De même, par continuité du passage à l'inverse, la suite (B_p^{-1}) converge vers B^{-1}, car B est inversible, donc elle est bornée pour $\|\cdot\|_2$.

Or, pour $M \in \mathscr{S}_n^{++}(\mathbb{R})$, on a :

$$\begin{aligned}\|M\|_2 &= \sqrt{\rho({}^tMM)} \text{ (corollaire 1.6)}\\ &= \sqrt{\rho(M^2)}\\ &= \rho(M) = \max\{\mathrm{Spec}(M)\}.\end{aligned}$$

Donc, tous les spectres des B_p sont majorés par une constante C. En appliquant le même raisonnement à B_p^{-1}, on voit que les spectres des B_p sont minorés par une constante C'. Par suite, toutes les valeurs propres des B_p sont contenues dans le compact $K = [C', C] \subset]0, +\infty[$, puis les valeurs propres des A_p, qui sont les logarithmes des valeurs propres des B_p, sont dans le compact $[\ln C', \ln C]$ de $\mathbb{R}$.

Ainsi, la suite (A_p) est bornée pour $\|\cdot\|_2$. Or, sa seule valeur d'adhérence est A. Soit en effet (A_{p_k}) une sous-suite qui converge vers $\overline{A}$. Du fait que l'on a : $\exp A_{p_k} = B_{p_k}$ pour tout k, on en déduit : $\exp \overline{A} = B = \exp A$. Par injectivité de l'exponentielle sur $\mathscr{S}_n^{+}(\mathbb{R})$, il vient : $\overline{A} = A$. Comme dans un espace vectoriel de dimension finie, une suite bornée qui n'a qu'une valeur d'adhérence converge vers cette valeur d'adhérence, on peut conclure.

(ii) Similaire. Laissé en exercice □

Avec l'aide de l'exponentielle, on en arrive aux conséquences topologiques annoncées du théorème de décomposition polaire.

2.4. Corollaire. *On a des homéomorphismes :*

$$\mathrm{GL}_n(\mathbb{R}) \cong \mathrm{O}_n(\mathbb{R}) \times \mathbb{R}^{n(n+1)/2} \qquad \textit{et} \qquad \mathrm{GL}_n(\mathbb{C}) \cong \mathrm{U}_n(\mathbb{C}) \times \mathbb{R}^{n^2}.$$

2.5. Questions d'injectivité et surjectivité de l'exponentielle

Notons que l'exponentielle, définie de l'espace de toutes les matrices vers le groupe linéaire, n'est pas injective. Sur $\mathbb{C}$, elle ne l'est pas dès la dimension 1 : $\exp(2ik\pi) = 1$ pour tout entier k. Sur $\mathbb{R}$, c'est à partir de la dimension 2 :

$$\exp\begin{pmatrix} 0 & -2k\pi \\ 2k\pi & 0 \end{pmatrix} = \mathrm{I}_2, \quad (k \in \mathbb{Z}).$$

En fait, la propriété 2.1 (vii) de compatibilité de l'exponentielle avec le spectre[2] permet d'étudier en profondeur ces problèmes d'injectivité.

Du côté de la surjectivité, l'exponentielle complexe est surjective : on le voit facilement à l'aide de la décomposition de Dunford multiplicative. L'image de l'exponentielle réelle est égale à l'ensemble des carrés de matrices réelles, voir l'exercice B.12. On se contente ici de vérifier le défaut de surjectivité.

2.5.1. Exemple. On a :

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \exp\begin{pmatrix} 0 & \pi \\ -\pi & 0 \end{pmatrix}.$$

En revanche, la matrice

$$\begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}$$

n'est pas l'exponentielle d'une matrice réelle. En effet, si A était une telle matrice, son spectre contiendrait deux complexes λ et μ tels que $e^{\lambda} = -1$, $e^{\mu} = -2$, mais aussi leur complexes conjugués, qui sont distincts. Ce qui pose un problème vu que A n'a au plus que deux valeurs propres.

A. Développement. Application de la décomposition polaire à l'étude de $\mathrm{O}(p, q)$

A.1. Définition. Soit p et q deux entiers naturels. On note $\mathrm{O}(p, q)$ le sous-groupe de $\mathrm{GL}_{p+q}(\mathbb{R})$ formé des isométries de la forme quadratique standard

[2] Cette propriété se généralise à d'autres fonctions analytiques.

sur $\mathbb{R}^{p+q}$ de signature (p,q), c'est-à-dire :

$$x_1^2+\cdots+x_p^2-x_{p+1}^2-\cdots-x_{p+q}^2,$$

dont la matrice dans la base canonique est :

$$\mathrm{I}_{(p,q)}=\mathrm{diag}(\underbrace{1,\ldots,1}_{p},\underbrace{-1,\ldots,-1}_{q}).$$

Par abus, on note aussi parfois $\mathrm{O}(p)$ le groupe $\mathrm{O}(p,\mathbb{R})=\mathrm{O}_p(\mathbb{R})$. (Attention, au contraire de $\mathrm{O}(p,q,\mathbb{C})$, qui n'a pas de sens, il y a un risque de confusion avec $\mathrm{O}(p,\mathbb{C})$.)

Nous avons choisi de présenter une application de la décomposition polaire à l'étude des groupes $\mathrm{O}(p,q)$. Il va en résulter une description de sa composante neutre. Contrairement au cas de $\mathrm{O}(n,\mathbb{R})$, il ne s'agira pas de $\mathrm{SL}_n(\mathbb{R})\cap\mathrm{O}(p,q)$, mais d'un sous-groupe d'indice 2 de cette intersection, appelé *groupe orthochrone*, nom qui lui va comme un gant, mais qui irait aussi à un groupe de *heavy metal* à condition qu'il respecte le tempo.

A.2. Proposition. *Soit $p,q\neq 0$. Il existe un homéomorphisme :*

$$\mathrm{O}(p,q)\cong\mathrm{O}(p)\times\mathrm{O}(q)\times\mathbb{R}^{pq}.$$

Démonstration. Soit $M\in\mathrm{O}(p,q)\subset\mathrm{GL}_n(\mathbb{R})$, avec $n=p+q$. Par décomposition polaire, il existe deux matrices $O\in\mathrm{O}(n,\mathbb{R})$ et $S\in\mathscr{S}_n^{++}(\mathbb{R})$ telles que $M=OS$. On veut montrer que S et O sont dans $\mathrm{O}(p,q)$. Pour cela, il suffit de montrer que S l'est.

Soit $T={}^tMM$. On a comme à l'accoutumée : $S^2=T$.

Montrons d'abord que $\mathrm{O}(p,q)$ est stable par transposition :

$$\begin{aligned}M\in\mathrm{O}(p,q)&\Longrightarrow M\mathrm{I}_{(p,q)}\,{}^tM=\mathrm{I}_{(p,q)}\\&\Longrightarrow {}^tM^{-1}\mathrm{I}_{(p,q)}M^{-1}=\mathrm{I}_{(p,q)}\\&\Longrightarrow {}^tM^{-1}\in\mathrm{O}(p,q)\\&\Longrightarrow {}^tM\in\mathrm{O}(p,q).\end{aligned}$$

On en déduit : $T={}^tMM\in\mathrm{O}(p,q)$.

On a donc $S^2\in\mathrm{O}(p,q)$. Pour conclure sur S, il va falloir prendre la « racine carrée » de T, donc, via l'exponentielle, diviser par deux. Comme T est, comme S, définie positive, on peut écrire $T=\exp U$ pour $U\in\mathscr{S}_n(\mathbb{R})$

convenable. On a alors :

$$\begin{aligned}
T \in \mathrm{O}(p,q) &\iff T\mathrm{I}_{(p,q)}\,{}^tT = \mathrm{I}_{(p,q)} \\
&\iff {}^tT = \mathrm{I}_{(p,q)}T^{-1}\mathrm{I}_{(p,q)}^{-1} \\
&\iff {}^t\exp U = \mathrm{I}_{(p,q)}(\exp U)^{-1}\mathrm{I}_{(p,q)}^{-1} \\
&\iff \exp {}^tU = \mathrm{I}_{(p,q)}\exp(-U)\,\mathrm{I}_{(p,q)}^{-1} \\
&\iff \exp {}^tU = \exp(-\mathrm{I}_{(p,q)}U\mathrm{I}_{(p,q)}^{-1}) \\
&\iff {}^tU = U = -\mathrm{I}_{(p,q)}U\mathrm{I}_{(p,q)}^{-1} \ (\exp : \mathscr{S}_n \to \mathscr{S}_n^{++} \text{ bijective}) \\
&\iff U\mathrm{I}_{(p,q)} + \mathrm{I}_{(p,q)}U = 0 \qquad \text{(condition linéaire)} \\
&\iff \frac{U}{2}\mathrm{I}_{(p,q)} + \mathrm{I}_{(p,q)}\frac{U}{2} = 0 \\
&\iff \frac{{}^tU}{2} = -\mathrm{I}_{(p,q)}\frac{U}{2}\mathrm{I}_{(p,q)}^{-1} \\
&\iff \exp\Big(\frac{{}^tU}{2}\Big) = \exp\Big(-\mathrm{I}_{(p,q)}\frac{U}{2}\mathrm{I}_{(p,q)}^{-1}\Big) \\
&\iff {}^t\exp\Big(\frac{U}{2}\Big) = \mathrm{I}_{(p,q)}\Big(\exp\big(\frac{U}{2}\big)\Big)^{-1}\mathrm{I}_{(p,q)}^{-1}.
\end{aligned}$$

Or, on a : $\exp(U/2) \in \mathscr{S}_n(\mathbb{R})$ et $\exp^2(U/2) = \exp U = T$. On en déduit alors : $\exp(U/2) = S$ et $S\mathrm{I}_{(p,q)}\,{}^tS = \mathrm{I}_{(p,q)}$, c'est-à-dire : $S \in \mathrm{O}(p,q)$. Et voilà : $O \in \mathrm{O}(p,q)$, et la décomposition polaire $M = OS \mapsto (O,S)$ induit une bijection bicontinue :

$$\mathrm{O}(p,q) \cong \big(\mathrm{O}(p,q) \cap \mathrm{O}(n)\big) \times \big(\mathrm{O}(p,q) \cap \mathscr{S}_n^{++}(\mathbb{R})\big).$$

Reste à comprendre le groupe $\mathrm{O}(p,q) \cap \mathrm{O}(n)$ et l'intersection $\mathrm{O}(p,q) \cap \mathscr{S}_n^{++}$.

Pour le premier, soit $O \in \mathrm{O}(p,q) \cap \mathrm{O}(n)$. On découpe O en blocs, il vient :

$$O = \begin{pmatrix} A & C \\ B & D \end{pmatrix} \in \mathrm{O}(p,q) \iff \begin{cases} {}^tAA - {}^tBB = \mathrm{I}_p \\ {}^tAC - {}^tBD = 0 \\ {}^tCA - {}^tDB = 0 \\ {}^tCC - {}^tDD = -\mathrm{I}_q, \end{cases}$$

car

$$\begin{aligned}
\begin{pmatrix} \mathrm{I}_p & 0 \\ 0 & -\mathrm{I}_q \end{pmatrix} &= \begin{pmatrix} {}^tA & {}^tB \\ {}^tC & {}^tD \end{pmatrix}\begin{pmatrix} \mathrm{I}_p & 0 \\ 0 & -\mathrm{I}_q \end{pmatrix}\begin{pmatrix} A & C \\ B & D \end{pmatrix} \\
&= \begin{pmatrix} {}^tA & {}^tB \\ {}^tC & {}^tD \end{pmatrix}\begin{pmatrix} A & C \\ -B & -D \end{pmatrix} = \begin{pmatrix} {}^tAA - {}^tBB & {}^tAC - {}^tBD \\ {}^tCA - {}^tDB & {}^tCC - {}^tDD \end{pmatrix}.
\end{aligned}$$

D'autre part, on a :

$$O \in \mathrm{O}(n) \Longleftrightarrow \begin{cases} {}^tAA + {}^tBB = \mathrm{I}_p \\ {}^tAC + {}^tBD = 0 \\ {}^tCA + {}^tDB = 0 \\ {}^tCC + {}^tDD = \mathrm{I}_q, \end{cases}$$

car

$$\begin{pmatrix} \mathrm{I}_p & 0 \\ 0 & \mathrm{I}_q \end{pmatrix} = \begin{pmatrix} {}^tA & {}^tB \\ {}^tC & {}^tD \end{pmatrix} \begin{pmatrix} A & C \\ B & D \end{pmatrix} = \begin{pmatrix} {}^tAA + {}^tBB & {}^tAC + {}^tBD \\ {}^tCA + {}^tDB & {}^tCC + {}^tDD \end{pmatrix}.$$

De ${}^tBB = 0$, on tire : $\mathrm{tr}({}^tBB) = 0$. Si $B = (b_{ij})$, il vient : $\sum_{i,j} b_{ij}^2 = 0$, puis : $B = 0$. De même, C est la matrice nulle et, par suite : $A \in \mathrm{O}(p)$ et $D \in \mathrm{O}(q)$. Ainsi :

$$\mathrm{O}(p,q) \cap \mathrm{O}(n) = \left\{ \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} \ , \ A \in \mathrm{O}(p), D \in \mathrm{O}(q) \right\} \cong \mathrm{O}(p) \times \mathrm{O}(q).$$

Pour la seconde intersection, on utilise à nouveau l'exponentielle : d'une part, $\exp : \mathscr{S}_n(\mathbb{R}) \to \mathscr{S}_n^{++}(\mathbb{R})$ est un homéomorphisme ; de même, comme on l'a vu plus haut, que :

$$\exp : L = \left\{ U \in \mathscr{M}_n(\mathbb{R}) \ , \ U\mathrm{I}_{(p,q)} + \mathrm{I}_{(p,q)}U = 0 \right\} \to \mathrm{O}(p,q),$$

d'où l'homéomorphisme :

$$\mathscr{S}_n(\mathbb{R}) \cap L \cong \mathrm{O}(p,q) \cap \mathscr{S}_n^{++}(\mathbb{R}).$$

On sait que $\mathscr{S}_n(\mathbb{R})$ est un espace vectoriel réel de dimension $n(n+1)/2$, et l'on montre en exercice que son intersection avec L est de dimension pq.

On obtient finalement : $\mathrm{O}(p,q) \cap \mathscr{S}_n^{++}(\mathbb{R}) \cong \mathbb{R}^{pq}$, d'où l'homéomorphisme :

$$\mathrm{O}(p,q) \cong \mathrm{O}(p) \times \mathrm{O}(q) \times \mathbb{R}^{pq}.$$

□

A.3. Remarque. Cette preuve mérite d'être photographiée, car elle utilise une méthode applicable à beaucoup de groupes. En particulier, même si cela peut paraître un détail, la stabilité de sous-groupes par la transposition demande un peu d'attention (voir l'exercice B.4).

A.4. Corollaire. *Soit p et q deux entiers naturels. Le groupe $\mathrm{O}(p,q)$ est compact si et seulement si $p = 0$ ou $q = 0$.*

A.5. Corollaire. *Soit p et q deux entiers naturels non nuls.*

(i) *Le groupe topologique $\mathrm{O}(p,q)$ admet quatre composantes connexes,*

(ii) *La composante connexe de l'identité est :*

$$\mathrm{SO}_0(p,q) = \left\{ \begin{pmatrix} A & C \\ B & D \end{pmatrix} \in \mathrm{SO}(p,q) \ : \ A \in \mathrm{GL}_p^+(\mathbb{R}) \right\}.$$

(iii) *On a :* $\mathrm{O}(p,q) \simeq \mathrm{SO}_0(p,q) \rtimes (\mathbb{Z}/2\mathbb{Z})^2$.

A.6. Remarque. Par contraste, la composante connexe de l'identité dans $\mathrm{O}(n,\mathbb{K})$ est $\mathrm{SO}(n,\mathbb{K})$, qui est d'indice 2 dans $\mathrm{O}(n,\mathbb{K})$, où $\mathbb{K}$ désigne le corps complexe ou réel.

Démonstration. (i) Vu l'homéomorphisme de A.2 et vu que $\mathrm{O}(n)$ a deux composantes connexes, le groupe $\mathrm{O}(p,q)$ en a quatre.

(ii) Le déterminant envoie $\mathrm{O}(p,q)$ sur $\{\pm 1\}$: par exemple, si l'on pose :

$$g = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix},$$

alors $\det(g) = -1$. Comme son noyau est $\mathrm{SO}(p,q)$, on a une suite exacte :

$$1 \longrightarrow \mathrm{SO}(p,q) \longrightarrow \mathrm{O}(p,q) \xrightarrow{\det} \{\pm 1\} \longrightarrow 1.$$

Cela exprime que : $\mathrm{O}(p,q)/\mathrm{SO}(p,q) \simeq \{\pm 1\}$. Ainsi, $\mathrm{O}(p,q)$ est la réunion disjointe de $\mathrm{SO}(p,q)$ et $g\,\mathrm{SO}(p,q)$. Par suite, $\mathrm{SO}(p,q)$ a deux composantes connexes. Soit $M \in \mathrm{SO}(p,q)$, écrite par blocs ainsi :

$$M = \begin{pmatrix} A & C \\ B & D \end{pmatrix}.$$

On a vu que ${}^tAA = \mathrm{I}_n + {}^tCC$, donc ${}^tAA \in \mathscr{S}_n^{++}(\mathbb{R})$, car $\langle Ax, Ax\rangle = \langle x,x\rangle + \langle Cx, Cx\rangle > 0$, pour $x \neq 0$. En particulier, A est inversible, ce qui permet d'associer, à M comme ci-dessus, le signe $\det(A)/|\det(A)|$. On définit ainsi un morphisme continu (fraction de polynômes) $f : \mathrm{SO}(p,q) \to \{\pm 1\}$.

Le morphisme f est surjectif, car en prenant

$$h = \begin{pmatrix} -1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & -1 \end{pmatrix},$$

on a : $f(h) = -1$. Par continuité, f est constante sur la composante connexe de l'identité, si bien que l'on a : $f|_{\mathrm{SO}_0(p,q)} = 1$, puis : $f|_{h\,\mathrm{SO}_0(p,q)} = -1$. Finalement, les composantes connexes de $\mathrm{SO}(p,q)$ sont $\mathrm{SO}_0(p,q)$ et $h\,\mathrm{SO}_0(p,q)$. Enfin, on a bien :

$$\mathrm{SO}_0(p,q) = \mathrm{Ker}(f) = \left\{ \begin{pmatrix} A & C \\ B & D \end{pmatrix} \in \mathrm{SO}(p,q) \,:\, A \in \mathrm{GL}_p^+(\mathbb{R}) \right\}.$$

(iii) Laissé en exercice. On peut prendre pour système de représentants de $\mathrm{O}(p,q)/\,\mathrm{SO}_0(p,q)$ l'ensemble $H = \{\mathrm{I}_{p+q}, g, h, gh\}$ et remarquer que H est un sous-groupe de $\mathrm{O}(p,q)$ isomorphe à $(\mathbb{Z}/2\mathbb{Z})^2$. □

A.7. Remarque. On a vu que $\mathrm{O}(p,q)$ agissait de façon transitive sur les nappes de $\mathbb{R}^{p+q}$ d'équation

$$\sum_{i=1}^{p} x_i^2 - \sum_{i=p+1}^{p+q} x_i^2 = c \quad (c \in \mathbb{R} \text{ fixé}).$$

On montre facilement qu'il en est de même de $\mathrm{SO}(p,q)$ en dimension 1 ou 2. En conclusion, une telle nappe a au plus deux composantes connexes. En dimension 3, il existe des hyperboloïdes à une nappe et des hyperboloïdes à deux nappes. Par le théorème de Sylvester, on voit que les composantes connexes possibles pour une quadrique sont au nombre de 1 ou 2. Lorsqu'il y a deux composantes connexes, le groupe $\mathrm{SO}_0(p,q)$ agit transitivement sur chacune des composantes connexes, mais pas sur toute la nappe, par connexité. L'exercice B.8 montre que l'on a deux composantes connexes si et seulement si la signature de la forme quadratique est $(1,v)$ ou $(u,1)$.

A.8. Remarque. Le groupe $\mathrm{SO}_0(p,q)$ est donc appelé *groupe orthochrone* en lien avec la forme de Lorentz $\sum_{j=1}^{p} x_i^2 - \sum_{k=1}^{q} x_{p+k}^2$ (c'est-à-dire $x^2 + y^2 + z^2 - t^2$ si $(p,q) = (3,1)$). Le groupe orthochrone agit sur ses nappes (les lignes de niveau) en préservant l'orientation du temps, ce qui explique son nom.

Pour vérifier si vous avez bien compris les méthodes de ce chapitre, vous pouvez faire l'exercice suivant.

A.9. Exercice (Décomposition polaire hermitienne)

Soit $n \in \mathbb{N}^*$. Exhiber un homéomorphisme : $\mathrm{O}_n(\mathbb{C}) \cong \mathrm{O}_n(\mathbb{R}) \times \mathbb{R}^{n(n-1)/2}$.

Référence. L'ouvrage [25], chapitres 1, 3, 4 et 5.

B. Exercices

Décompositions polaires de groupes classiques et applications

B.1. Exercice (La décomposition polaire de $\mathrm{GL}_n(\mathbb{C})$)

1. *Montrer que, pour tout* $A \in \mathrm{GL}_n(\mathbb{C})$, *il existe une unique* $H \in \mathscr{H}_n^{++}$ *telle que* $H^2 = A^*A$.
   Montrer que $U := AH^{-1} \in U(n)$.
2. *En déduire que, pour tout* $A \in \mathrm{GL}_n(\mathbb{C})$, *il existe unique* $H \in \mathscr{H}_n^{++}$ *et* $U \in U(n)$ *telles que* $A = UH$.
3. *Montrer que la réciproque de la multiplication* $\mathrm{U}(n) \times \mathscr{H}_n^{++} \longrightarrow \mathrm{GL}_n(\mathbb{C})$ *est continue.*
4. *En déduire que* $\mathrm{GL}_n(\mathbb{C})$ *est homéomorphe à* $\mathrm{U}(n) \times \mathbb{R}^{n^2}$.

B.2. Exercice (Décomposition polaire de $\mathrm{O}(n, \mathbb{C})$)

1. *Montrer que si* $U \in \mathrm{U}(n)$ *satisfait à* ${}^tUU = \mathrm{I}_n$ *alors* $U \in \mathrm{O}(n) = \mathrm{O}(n, \mathbb{R})$.
2. *Soit* $M \in \mathrm{O}(n, \mathbb{C})$ *et* $M = UH$ *la décomposition polaire vue comme un élément de* $\mathrm{GL}_n(\mathbb{C})$, i.e. $U \in \mathrm{U}(n)$ *et* $H \in \mathscr{H}_n^{++}$. *Montrer la relation :* ${}^tH({}^tUU) = H^{-1}$.
3. *En déduire que* $U \in \mathrm{O}(n)$ *et que* H *s'écrit sous la forme* $\exp(iA)$, *où* A *est une matrice anti-symétrique réelle.*
4. *En déduire que* $\mathrm{O}(n, \mathbb{C})$ *est homéomorphe à* $\mathrm{O}(n) \times \mathbb{R}^{\frac{1}{2}n(n-1)}$.

B.3. Exercice (Décomposition polaire de $\mathrm{U}(p, q)$)

L'espace $\mathbb{C}^n$ *est muni du produit scalaire hermitien standard :*

$$\langle X, Y \rangle := X^*Y, \qquad X, Y \in \mathbb{C}^n.$$

On désigne par A^* *la transposée de la conjuguée de* A. *On considère les ensembles :*

$$\mathrm{U}(n) := \{A \in \mathscr{M}_n(\mathbb{C}) | A^*A = 1\},\ \mathrm{U}(p,q) := \{A \in \mathscr{M}_n(\mathbb{C}) | A^*\mathrm{I}_{(p,q)}A = \mathrm{I}_{(p,q)}\},$$

$$\mathscr{H}_n := \{A \in \mathscr{M}_n(\mathbb{C}) | A^* = A\},\ \mathscr{H}_n^{++} := \{A^*A | A \in \mathrm{GL}_n(\mathbb{C})\}.$$

Soit enfin : $\mathrm{I}_{(p,q)} = \mathrm{diag}(\underbrace{1, \cdots, 1}_{p}, \underbrace{-1, \cdots, -1}_{q})$.

1.(a) *Montrer que* $A \mapsto \mathrm{tr}(A^*A)$ *définit une norme de l'espace complexe* $\mathscr{M}_n(\mathbb{C})$.
   (b) *Montrer que* $\mathrm{U}(n)$ *est un sous-groupe compact de* $\mathrm{GL}_n(\mathbb{C})$.

2. *Montrer que les groupes topologiques* $\mathrm{U}(p,q) \cap U(p+q)$ *et* $\mathrm{U}(p) \times U(q)$ *sont isomorphes.*
3. *Soit* S *une matrice de* $\mathscr{H}_n^{++}$.
   (a) Montrer qu'il existe une matrice R *de* $\mathscr{H}_n^{++}$ *telle que* $R^2 = S$.
   (b) Montrer qu'une telle matrice R *est alors unique.*
4. *Montrer que les conditions suivantes sont équivalentes :*
   (a) $R \in \mathrm{U}(p,q) \cap \mathscr{H}_n^{++}$,
   (b) $R = \exp(T)$ *avec* $T \in \mathscr{H}_{p,q} := \{M \in \mathscr{M}_n(\mathbb{C}) | M^* \mathrm{I}_{(p,q)} + \mathrm{I}_{(p,q)} M = 0\}$.
5. *Soit* $A = VR$, *avec* $(V,R) \in \mathrm{U}(p+q) \times \mathscr{H}_{p+q}^{++}$, *la décomposition polaire hermitienne de* $A \in \mathrm{GL}_{p+q}(\mathbb{C})$. *Montrer que les conditions suivantes sont équivalentes :*
   (a) $A \in \mathrm{U}(p,q)$ *;*
   (b) $R \in \mathrm{U}(p,q)$ *et* $V \in \mathrm{U}(p,q)$.
6. *Montrer que l'application*

$$\begin{aligned} \psi : \mathrm{U}(p) \times \mathrm{U}(q) \times \mathscr{H}_{p,q} &\longrightarrow \mathrm{U}(p,q) \\ (V_p, V_q, T) &\longmapsto \begin{pmatrix} V_p & 0 \\ 0 & V_q \end{pmatrix} \exp(T) \end{aligned}$$

   est un homéomorphisme.
7. *Calculer* $\dim \mathscr{H}_{p,q}$. *Est-ce que* $\mathrm{U}(p,q)$ *est compact ?*
8. *Montrer que* $\mathrm{U}(p)$ *est connexe.*
   On pourra admettre qu'une sphère est connexe en toute dimension strictement positive.
9. *Montrer que* $\mathrm{U}(p,q)$ *est connexe.*

B.4. Exercice. *Montrer que les groupes matriciels suivants sont stables par la transposition :* $\mathrm{U}(n,\mathbb{C})$, $\mathrm{O}(n,\mathbb{K})$, $\mathrm{O}(p,q)$, *où* $\mathbb{K}$ *est un corps quelconque et* n, p, q *sont des entiers naturels non nuls.*

On pourra jeter un coup d'œil à la preuve de la proposition A.2.

B.5. Exercice (Adhérences des classes de congruences de matrices symétriques réelles)

Soit n *un entier naturel non nul. On rappelle que la signature d'une matrice symétrique est le couple* (p,q) *où* p *(resp.* q*) est le nombre de valeurs propres strictement positives (resp. strictement négatives) de la matrice. Étant donné deux entiers* p *et* q, *on note* $\mathscr{O}_{p,q}$ *la partie de* $\mathscr{S}_n$ *formée des matrices de signature* (p,q) *: ce sont les orbites pour l'action par congruence de* $\mathrm{GL}_n(\mathbb{R})$ *sur* $\mathscr{S}_n = \mathscr{S}_n(\mathbb{R})$.

1. *On note* GL_n^+ *le sous-groupe du groupe des matrices réelles* $n\times n$ *inversibles de déterminant strictement positif.*

   (a) Démontrer rapidement que GL_n^+ *est connexe.*

   On pourra utiliser la décomposition polaire ou la connexité du groupe des matrices de déterminant 1.

   (b) On fait agir le groupe GL_n^+ *sur* $\mathscr{S}_n$ *par congruence : pour* $g \in \mathrm{GL}_n^+$ *et* $A \in \mathscr{S}_n$*, l'action de* g *sur* A *est donnée par :* $g \cdot A = g\,A\,{}^t g$*. Montrer que les orbites de cette action sont les parties* $\mathscr{O}_{p,q}$*, où* (p,q) *est un couple d'entiers naturels tels que* $p+q \leqslant n$*.*

   (c) En déduire que pour tout couple (p,q)*, l'orbite* $\mathscr{O}_{p,q}$ *est connexe.*

2. *(a) Dans cette question, on suppose que* $n = 2$*. Montrer que* $\mathscr{O}_{1,1}$ *est ouverte, mais pas fermée.*

   (b) Dans cette question, n *est quelconque. Démontrer que* $\mathscr{O}_{p,q}$ *est ouverte si et seulement si* $p+q=n$*. Démontrer que* $\mathscr{O}_{p,q}$ *est fermée si et seulement si* $p=q=0$*.*

   On va généraliser ces constatations.

3. *Montrer que l'application* $p : \mathscr{S}_n \to \{0,\dots,n\}$*,* $A \mapsto p(A)$*, qui associe à* A *la première composante* $p(A)$ *de sa signature, est semi-continue inférieurement.*

   On pourra réaliser que $p(A)$ est le nombre de valeurs propres strictement positives de A et utiliser par exemple l'exercice II-F.33.

4. *Montrer que l'adhérence de* $\mathscr{O}_{p,q}$*, pour* $p+q=n$*, est donnée par :*

$$\overline{\mathscr{O}_{p,q}} = \bigsqcup_{\substack{0\leqslant h\leqslant p\\ 0\leqslant h'\leqslant q}} \mathscr{O}_{p-h,q-h'}.$$

5. *Soit* $\mathscr{S}_n^\times$ *l'ensemble des matrices symétriques inversibles. Montrer que les composantes connexes de* $\mathscr{S}_n^\times$ *sont les* $\mathscr{O}_{p,q}$*, où* (p,q) *parcourt l'ensemble des couples tels que* $p+q=n$*.*

Applications de la décomposition polaire

B.6. Exercice. *Soit* q *une forme quadratique non dégénérée* q *sur* $\mathbb{R}$*. Montrer que le groupe orthogonal* $\mathrm{O}(q)$ *est compact si et seulement si* q *est définie.*

B.7. Exercice (Des matrices réelles U(n)-semblables sont O(n)-semblables.). *Soit* $\mathscr{H}_n^+$ *l'ensemble des matrices hermitiennes positives (pas nécessairement définies) de* $\mathscr{M}_n(\mathbb{C})$*. Soit* $Q \in \mathscr{H}_n^+$*, que l'on diagonalise dans une base orthonormée :* $Q = PDP^{-1}$*, où* $P \in \mathrm{U}(n)$ *et* $D = \mathrm{diag}(d_1,\cdots,d_n)$ *avec* $d_i \in \mathbb{R}_+$*.*

1. *Soit* $R := P\sqrt{D}P^{-1}$. *Montrer que* $R^2 = Q$ *et il existe un polynôme* $q \in \mathbb{R}[X]$ *tel que* $R = q(Q)$.
2. *En déduire que si* $A \in \mathscr{M}_n(\mathbb{C})$ *commute avec* Q, *alors elle commute avec* R.
3. *Soit* $A, B \in \mathscr{M}_n(\mathbb{R})$ *et* $U \in U(n)$ *telles que* $A = UBU^*$.
   (a) *Soit* $X, Y \in \mathscr{M}_n(\mathbb{R})$ *telles que* $X + iY = U$. *Montrer qu'il existe* $\eta \in \mathbb{R}$ *tel que* $X + \eta Y \in \mathrm{GL}_n(\mathbb{R})$. *On pose* $P := X + \eta Y$.
   (b) *Montrer que* $A = PBP^{-1}$ *et* ${}^tA = P\,{}^tBP^{-1}$. *En déduire que* B *commute avec* ${}^tPP$.
   (c) *Soit* $P = VR$ *la décomposition polaire avec* $V \in \mathrm{O}(n)$ *et* $R \in \mathscr{S}_n^{++}$. *Montrer que* B *commute avec* R. *En déduire que* A *et* B *sont* $\mathrm{O}(n)$*-semblables.*

B.8. Exercice (Composantes connexes des quadriques réelles)

Soit q *une forme quadratique réelle non nulle de signature* (r, s). *On veut montrer que le nombre* k *de composantes connexes de la quadrique* $q(x) = a$, $a \neq 0$ *est donné par*

$$k = \begin{cases} 0 & \text{si } (r = 0 \text{ et } a < 0) \text{ ou } (s = 0 \text{ et } a > 0) \\ 1 & \text{si } (r \geqslant 2 \text{ et } a > 0) \text{ ou } (s \geqslant 2 \text{ et } a > 0) \\ 2 & \text{si } (r = 1 \text{ et } a > 0) \text{ ou } (s = 1 \text{ et } a < 0). \end{cases}$$

1. *On suppose* $n \leqslant m$, $r \geqslant 2$ *et l'on commence par étudier la quadrique* $\mathscr{C}$:

$$\mathscr{C} := \{(x_i) \in \mathbb{R}^m,\ x_1^2 + x_2^2 + \cdots + x_r^2 - x_{r+1}^2 - \cdots - x_n^2 = 1\}.$$

   (a) *Montrer que* $(1, 0, \ldots, 0)$ *et* $(-1, 0, \ldots, 0)$ *sont dans la même composante connexe de* $\mathscr{C}$.

   Regarder l'intersection du plan $\{(x_1, x_2, 0, \ldots, 0),\ (x_1, x_2)^2 \in \mathbb{R}^2\}$ avec $\mathscr{C}$.

   (b) *Montrer que* $\mathrm{O}(2,2)$ *agit transitivement sur* $\mathscr{C}$ *et que le quotient* $\mathrm{O}(2,2)/\mathrm{SO}_0(2,2)$ *agit sur les composantes connexes de* $\mathscr{C}$.

   On pourra considérer l'action de $\mathrm{O}(2,2)$ sur l'ensemble des composantes connexes de $\mathscr{C}$, puis remarquer que pour une action continue, un groupe connexe (tout comme son élément neutre) stabilise chaque composante connexe de l'espace topologique sur lequel il agit.

   (c) *À l'aide de représentants explicites du quotient* $\mathrm{O}(2,2)/\mathrm{SO}_0(2,2)$ *dans* $\mathrm{O}(2,2)$, *montrer que* $\mathscr{C}$ *est connexe.*

2. *On suppose* $n \leqslant m$, *et l'on étudie* $\mathscr{C}'$:

$$\mathscr{C}' := \{(x_i) \in \mathbb{R}^m,\ x_1^2 - x_2^2 - \cdots - x_n^2 = 1\}.$$

(a) *Montrer comme dans 1.(c) que le nombre de composantes connexes de C' est inférieur ou égal à 2.*

(b) *Montrer que la projection sur la première composante de C' contient des éléments positifs, des éléments négatifs, mais ne contient pas 0. En déduire que C' n'est pas connexe.*

3. *Conclure.*

Exponentielle

B.9. Exercice (Restriction de l'exponentielle au cône nilpotent)

Montrer que l'exponentielle réalise un homéomorphisme entre l'ensemble des matrices nilpotentes et l'ensemble des matrices unipotentes.

Le logarithme défini à la proposition 2.1 est un polynôme, une fois restreint aux unipotentes...

B.10. Exercice. *Soit A une matrice complexe.*

1. *Montrer qu'il existe un polynôme P tel que $P(A) = \exp A$.*

   L'ensemble des polynômes en A (où A est fixé) est un sous-espace vectoriel de l'espace des matrices ; il est donc fermé.

2. *On suppose A inversible. Montrer qu'il existe un polynôme Q tel que $\exp Q(A) = A$.*

   On pourra utiliser la décomposition de Dunford et traiter séparément le cas diagonalisable et le cas nilpotent. On pourra se servir de l'exercice B.9.

B.11. Exercice. *Résoudre dans $\mathcal{M}_n(\mathbb{C})$ l'équation $\exp A = \mathrm{I}_n$.*

Utiliser la décomposition de Dunford additive et multiplicative. On pourra se servir de l'exercice B.9.

B.12. Exercice (Image de l'exponentielle)

1. *Montrer que si A est une matrice complexe diagonalisable inversible, alors il existe une matrice B telle que $\exp(B) = A$.*

   Utiliser une matrice de passage.

2. *Montrer que si A est une matrice complexe inversible, alors il existe une matrice B telle que $\exp(B) = A$. En déduire que l'exponentielle complexe est surjective.*

   Utiliser la décomposition de Dunford multiplicative de A, puis construire B par une décomposition additive.

3. *Montrer que si A est l'image* $\exp(B)$ *d'une matrice réelle B, alors A est le carré d'une matrice réelle.*

   Considérer $\exp(B/2)$.

4. *Montrer la réciproque : l'image de l'exponentielle réelle est égale à l'image de $B \mapsto B^2$.*

   Faire le cas diagonalisable en se rappelant de l'exercice III-D.1, puis le cas unipotent/nilpotent. Conclure par Dunford.

B.13. Exercice. *Pour x complexe, calculer* $\exp\begin{pmatrix} 0 & x \\ x & 0 \end{pmatrix}$.

Calculer les puissances paires et impaires de cette matrice par récurrence, puis se rappeler ses développements limités.

Chapitre VII

Le corps des quaternions

« Je préviens : on a la puissance de feu d'un croiseur et des flingues de concours. »

Les Tontons flingueurs, Michel Audiard, 1963.

Passant rapidement sur « la genèse du nombre chez l'enfant », on retient du lycée la tour de corps : $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$; on sait y insérer des corps intermédiaires, tels $\mathbb{Q}(\pi)$ ou la clôture algébrique de $\mathbb{Q}$. Cette belle tour s'arrête là puisque par le théorème de D'Alembert-Gauss, il n'existe aucun surcorps (commutatif) de degré fini sur $\mathbb{C}$.

Cependant, si l'on s'autorise le non-commutatif, on peut et on va construire un corps-non-commutatif[1] $\mathbb{H}$, appelé corps des quaternions, qui est de degré 2 sur $\mathbb{C}$. Après une définition relativement artificielle, nous en verrons deux descriptions qui le rendent plus naturel. Elles permettent une analogie entre le rôle pour les complexes du groupe des nombres complexes de module 1, qui s'identifie au groupe des rotations planes $\mathrm{SO}(2)$ ou au cercle unité $\mathbf{S}^1$, et le rôle pour les quaternions du groupe $\mathrm{SU}(2)$, qui s'identifie à la sphère $\mathbf{S}^3$.

Une façon efficace de motiver ce chapitre serait de dire que si l'on est convaincu que $\mathbb{C}$ est un bon objet pour faire de la géométrie plane (c'est clair, ça, non ?), $\mathbb{H}$ en est également un pour faire de la géométrie en dimension 3 ou 4. C'était là en fait le but de son auteur Sir Hamilton, qui s'est fourvoyé des années en cherchant un corps de dimension 3 sur $\mathbb{R}$, que nul ne trouva jamais, encouragé par son enfant qui lui demandait quotidiennement : *"Well, Papa, can you multiply triplets?"*[2].

[1] Il semble être devenu traditionnel d'appeler *corps* un corps commutatif et *corps-non-commutatif* ou *algèbre à division* un corps non commutatif.

[2] This is a true story. Voir l'article [26] de Sophie Morier-Genoud et Valentin Ovsienko.

Comme applications du corps des quaternions, on réalisera de façon agréable SU(2) (resp. SU(2) × SU(2)) comme revêtement universel (bien que cette notion ne soit pas définie ici) de SO(3) (resp. SO(4)).

1. Le corps des quaternions

1.1. Définition des quaternions

On lit sur une plaque à côté du *Broom Bridge* à Dublin : « le 16 octobre 1843, alors qu'il se promenait, Sir William Rowan Hamilton découvrit dans un éclair de génie la formule fondamentale sur la multiplication des quaternions [...] et la grava sur une pierre du pont ». La voici.

1.1.1. Définition. On appelle *corps des quaternions* et l'on note[3] $\mathbb{H}$ l'algèbre[4] de dimension 4 sur le corps des réels, ayant pour base $(\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k})$, dans laquelle la multiplication est définie par :

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -\mathbf{1} \quad \text{et} \quad \mathbf{ijk} = -\mathbf{1}.$$

Pour préciser, l'élément $\mathbf{1}$ est neutre et la dernière relation signifie[5] :

$$\mathbf{ij} = \mathbf{k} = -\mathbf{ji}, \quad \mathbf{jk} = \mathbf{i} = -\mathbf{kj}, \quad \mathbf{ki} = \mathbf{j} = -\mathbf{ik}.$$

1.1.2. Remarque. Pour l'instant, il n'est pas clair que l'algèbre $\mathbb{H}$ est un corps, ni même une algèbre associative. Cela sera montré d'ici peu dans le lemme 1.1.5 et le corollaire 1.1.6. On pourra aussi s'en convaincre de manière alternative à l'aide de la représentation matricielle des quaternions.

1.1.3. Exercice (Pour les lecteurs familiers des présentations). *Montrer que* $\mathbb{H}$ *est bien l'algèbre (associative unitaire) présentée par générateurs* $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ *et relations* $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1$, $\mathbf{ijk} = -1$.

La notion de présentation pour une algèbre est... introduite dans le livre [22].

On a besoin de quelques ingrédients supplémentaires pour manipuler les quaternions.

---

[3] En référence à Sir Hamilton, bien évidemment, mais dans ce chapitre, nous en profiterons pour rendre hommage à Haïti.

[4] *A priori*, il s'agit ici seulement d'un espace vectoriel muni d'une loi interne bilinéaire : de la sorte, la donnée du produit de deux vecteurs de base détermine le produit entier. On en verra l'associativité incontinent.

[5] Toute ressemblance avec le produit vectoriel dans $\mathbb{R}^3$ n'est certainement pas fortuite.

1.1.4. Définition. On identifie $\mathbb{R}$ à la sous-algèbre de $\mathbb{H}$ engendrée par $\mathbf{1}$. On note $\mathbb{I}$ le sous-espace de $\mathbb{H}$ suivant :

$$\mathbb{I} = \mathbb{R}\mathbf{i} \oplus \mathbb{R}\mathbf{j} \oplus \mathbb{R}\mathbf{k};$$

ses éléments sont appelés *imaginaires* ou *imaginaires quaternioniques.*

Pour $(x, y, z, t) \in \mathbb{R}^4$ et $h = x + y\mathbf{i} + z\mathbf{j} + t\mathbf{k} \in \mathbb{H}$, on appelle *conjugué* de h et l'on note $\overline{h}$ le quaternion :

$$\overline{h} = x - y\mathbf{i} - z\mathbf{j} - t\mathbf{k}.$$

On appelle *norme* de h le quaternion :

$$N(h) = h\overline{h}.$$

1.1.5. Lemme. *Soit $h, h', h'' \in \mathbb{H}$. On a :*

(i) $(hh')h'' = h(h'h'')$;

(ii) $h \in \mathbb{R} \iff \overline{h} = h$;

(iii) $h \in \mathbb{I} \iff h^2 \in \mathbb{R}^- \iff \overline{h} = -h$;

(iv) *si* $h = x + y\mathbf{i} + z\mathbf{j} + t\mathbf{k}$, $N(h) = h\overline{h} = \overline{h}h = x^2 + y^2 + z^2 + t^2 \in \mathbb{R}$;

(v) $N(hh') = N(h)N(h')$.

Démonstration. Laissée en exercice ; il ne s'agit que de calculs directs. Toutefois, l'égalité $N(hh') = N(h)N(h')$ prendra un relief nouveau une fois vue dans le paragraphe 1.2 la réalisation matricielle des quaternions, puisque la norme, dans cette réalisation, n'est autre que le déterminant. □

1.1.6. Corollaire. *L'algèbre $\mathbb{H}$ est un corps-non-commutatif.*

Démonstration. On a vu l'associativité dans le lemme ; il n'y a plus qu'à vérifier que tout élément non nul a un inverse : si $h = x + y\mathbf{i} + z\mathbf{j} + t\mathbf{k} \in \mathbb{H} \setminus \{0\}$, on a en effet :

$$h^{-1} = \frac{1}{N(h)}\overline{h}.$$

□

1.1.7. Remarque. Les lecteurs ayant passé le baccalauréat avant 2013 savent que $\mathbb{C}$ permet de résoudre algébriquement des problèmes de géométrie plane. De même, $\mathbb{H}$ est un outil algébrique pour travailler sur la géométrie en dimension 3 ou 4. Les quaternions (et leur multiplication) sont utilisées en pratique pour représenter les rotations de l'espace dans les ordinateurs et jeux vidéo.

La non-commutativité de l'algèbre des quaternions fait que l'on s'intéresse en premier à son centre. Comme **1** est central dans $\mathbb{H}$, il en est de même de la sous-algèbre $\mathbb{R}$. En fait, la réciproque est vraie.

1.1.8. Proposition. *Le centre de* $\mathbb{H}$, i.e. *l'ensemble des éléments de* $\mathbb{H}$ *qui commutent à tout* $\mathbb{H}$, *est réduit à* $\mathbb{R}$.

Démonstration. D'après l'assertion qui précède, il suffit de montrer une seule inclusion : soit h dans le centre de $\mathbb{H}$, montrons que h est réel.

Posons $h = x + y\mathbf{i} + z\mathbf{j} + t\mathbf{k}$, avec $x, y, z, t \in \mathbb{R}$. Alors, les égalités $h\mathbf{i} = \mathbf{i}h$, $h\mathbf{j} = \mathbf{j}h$, $h\mathbf{k} = \mathbf{k}h$ donnent facilement par identification $y = -y$, $z = -z$ et $t = -t$ et ainsi $h = x$ est réel. □

L'introduction de $\mathbb{H}$ comme nous venons de la faire n'est peut-être pas illuminante. Elle fournit un moyen pratique de faire des calculs. Mais pour comprendre la structure de $\mathbb{H}$, la raison de ses propriétés particulières (corps non-commutatif, norme, conjugaison...), nous allons présenter une description plus parlante. Pour procéder par analogie, faisons un petit saut en arrière et rappelons la construction de $\mathbb{C}$. En fait, cela ne fait pas de mal d'en rappeler deux.

1.2. Deux constructions de $\mathbb{C}$

Adjonction de racines. $\mathbb{C} := \mathbb{R}[X]/(X^2 + 1)$

Le polynôme X^2+1, de degré 2, est irréductible sur $\mathbb{R}$: il n'a pas de racine. Soit P un polynôme de $\mathbb{R}[X]$ non multiple de $X^2 + 1$, il est premier avec X^2+1. D'après la relation de Bézout, la classe de P modulo X^2+1 possède un inverse : il existe des polynômes U et V tels que :

$$UP + V(X^2 + 1) = 1.$$

Dans le quotient $\mathbb{R}[X]/(X^2+1)$ (on note avec barre au-dessus la classe dans le quotient), on obtient : $\overline{U}\,\overline{P} = \overline{1}$. Autrement dit, tout élément non nul de $\mathbb{R}[X]/(X^2 + 1)$ est inversible, et donc, l'anneau quotient est un corps.

Selon un procédé général[6], le corps des complexes ainsi obtenu est bien muni de l'automorphisme $\overline{\ }\ : z \mapsto \overline{z}$ et de la norme multiplicative $N(z) = z\overline{z}$ (produit des images de z par les deux seuls éléments du groupe de Galois).

[6] L'extension est galoisienne : séparable, car on est en caractéristique nulle, normale comme toute extension de degré 2.

Réalisation matricielle

On réalise $\mathbb{C}$ comme le cône engendré dans $\mathscr{M}_2(\mathbb{R})$ par le groupe des rotations :

$$\mathrm{SO}(2) = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in \mathscr{M}_2(\mathbb{R}) \,:\, (a,b) \in \mathbb{R}^2,\ a^2+b^2=1 \right\} \simeq \mathbf{S}^1.$$

Autrement dit :

$$\mathbb{C} \cong \{0\} \cup \mathbb{R}^{+*}\,\mathrm{SO}(2) = \mathbb{R}\,\mathrm{SO}(2) = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in \mathscr{M}_2(\mathbb{R}) \,:\, (a,b) \in \mathbb{R}^2 \right\}.$$

1.2.1. Remarque. Le cône $\mathbb{R}\,\mathrm{SO}(2)$ (attention, ce n'est pas un produit cartésien $\mathbb{R} \times \mathrm{SO}(2)$) est aussi un espace vectoriel, ce qui est n'est pas habituel. On peut vérifier directement que l'espace vectoriel $\mathbb{R}\,\mathrm{SO}(2)$ est même une sous-algèbre, qu'elle est commutative et que tout élément non nul est inversible.

L'intersection des sous-groupes $\mathbb{R}^{+*}$ et $\mathrm{SO}(2)$ de $\mathbb{C}^*$ est triviale. Comme le groupe $\mathbb{C}^*$ (abélien) est engendré par ces deux sous-groupes, il est isomorphe, par II-5.2.1, à leur produit (semi-)direct topologique.

On identifie de cette manière un nombre complexe $z = a+ib$, avec $a, b \in \mathbb{R}$, à une matrice

$$Z = a\mathbf{1} + b\mathbf{i} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad \text{où } \mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ et } \mathbf{i} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Dans cette réalisation, la conjugaison complexe est l'automorphisme de transposition : $Z \mapsto {}^tZ$ et la norme est le déterminant : $Z \mapsto N(Z) = \det Z$.

1.2.2. Remarque. Étant donné z et Z comme ci-dessus, Z est la matrice de l'application $m_z : \mathbb{C} \to \mathbb{C}$, $w \mapsto zw$ dans la base $(1, i)$ de l'espace vectoriel réel $\mathbb{C}$.

1.3. Réalisation matricielle complexe des quaternions

C'est la deuxième méthode qui va primer pour décrire les quaternions : $\mathbb{H}$ est le corps des « cumplexes », dans le sens qu'il suffit de remplacer les O par des U, et donc le groupe abélien $\mathrm{SO}(2)$ par $\mathrm{SU}(2)$, non abélien.

Figure 1.1. *Pas cun le mec !*

Considérons, dans $\mathrm{GL}_2(\mathbb{C})$, les sous-groupes $\mathbb{R}^{+*} = \mathbb{R}^{+*}\mathrm{I}_2$ et $\mathrm{SU}(2)$:

$$\mathrm{SU}(2) = \left\{ \begin{pmatrix} a & -\overline{b} \\ b & \overline{a} \end{pmatrix} \in \mathscr{M}_2(\mathbb{C}) \ : \ |a|^2 + |b|^2 = 1 \right\}.$$

Comme précédemment, leur intersection est triviale et ils commutent entre eux (dame ! $\mathbb{R}^{+*}$ est central dans $\mathrm{GL}_2(\mathbb{C})$), le groupe engendré par ces deux sous-groupes est isomorphe à leur produit (semi-)direct (topologique) : on le note H^* :

$$H^* \simeq \mathbb{R}^{+*} \times \mathrm{SU}(2) \quad \text{(produit direct de groupes)}.$$

On définit alors : $H = H^* \cup \{0\}$, de sorte que l'on a :

$$H = \left\{ \begin{pmatrix} a & -\overline{b} \\ b & \overline{a} \end{pmatrix} \in \mathscr{M}_2(\mathbb{C}) : \ a, b \in \mathbb{C} \right\}.$$

En tant qu'espace vectoriel réel, H est de dimension 4 et admet pour base :

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \mathbf{i} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \ \mathbf{j} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \ \mathbf{k} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Les trois dernières matrices sont appelées *matrices de Pauli.*

Ainsi, si $x, y, z, t \in \mathbb{R}$, si l'on note $a = x + iy$ et $b = -z + it$, un élément typique de H s'écrit :

$$h = x\mathbf{1} + y\mathbf{i} + z\mathbf{j} + t\mathbf{k} = \begin{pmatrix} a & -\overline{b} \\ b & \overline{a} \end{pmatrix}.$$

On vérifie que H est un corps-non-commutatif isomorphe à $\mathbb{H}$. Dans cette réalisation, l'anti-automorphisme de conjugaison $\overline{\ }$: $h \mapsto h^*$ s'identifie à l'adjonction des matrices et la norme multiplicative d'un élément h est : $N(h) = h\overline{h} = \overline{h}h = |a|^2 + |b|^2 = \det h$.

1.3.1. Remarque. Par analogie avec la remarque 1.2.2, on peut comprendre cette réalisation matricielle de $\mathbb{H}$ de la façon suivante. D'évidence, l'espace vectoriel $\mathbb{R}\mathbf{1} \oplus \mathbb{R}\mathbf{i}$ est une sous-algèbre réelle de $\mathbb{H}$ isomorphe à $\mathbb{C}$. On la notera $\mathbb{C}$ par abus. Comme $\mathbb{H}$ est une algèbre qui contient $\mathbb{C}$, c'est en particulier un espace vectoriel *complexe* de dimension 2. Quitte à être un peu redondant, précisons que le produit d'un complexe $a = u + iv$ (avec $u, v \in \mathbb{R}$) par un quaternion $h = x\mathbf{1} + y\mathbf{i} + z\mathbf{j} + t\mathbf{k}$ (avec $x, y, z, t \in \mathbb{R}$) est :

$$a \cdot h = (u\mathbf{1} + v\mathbf{i})(x\mathbf{1} + y\mathbf{i} + z\mathbf{j} + t\mathbf{k}) = \ldots$$

On va voir que les matrices qui engendrent H sont les matrices de la multiplication par $\mathbf{1}$, $\mathbf{i}$, $\mathbf{j}$ et $\mathbf{k}$ dans cet espace vectoriel complexe.

Attention ! La finesse, c'est que $\mathbb{C}$ n'est pas central dans l'algèbre $\mathbb{H}$: en effet, $\mathbf{i}$ ne commute pas à $\mathbf{j}$ et $\mathbf{k}$. Il en résulte que la multiplication *à gauche*

par un quaternion n'est pas une application linéaire de l'espace vectoriel complexe $\mathbb{H}$. On a donc recours à la multiplication *à droite*, qui, elle, est bien linéaire (sur les complexes).

Alors, pour $h = x\mathbf{1} + y\mathbf{i} + z\mathbf{j} + t\mathbf{k} \in \mathbb{H}$ comme ci-dessus, on peut vérifier que la matrice de l'application $m_h : \mathbb{H} \to \mathbb{H}$, $k \mapsto kh$ dans la base $(\mathbf{1}, \mathbf{j})$ de $\mathbb{H}$ est la matrice 2×2 associée à h. En symboles :

$$\begin{aligned} \mathbb{H} &\longrightarrow H \subset \mathscr{M}_2(\mathbb{C}) \\ h &\longmapsto \mathrm{mat}_{(\mathbf{1},\mathbf{j})}(m_h). \end{aligned}$$

1.4. Réalisation matricielle réelle des quaternions

Porté par l'élan consistant à remplacer un complexe par une matrice 2×2, on étend ce procédé aux matrices : on remplace chaque coefficient $a = x+iy$ (avec $x, y \in \mathbb{R}$) d'une matrice de H par la sous-matrice

$$\begin{pmatrix} x & -y \\ y & x \end{pmatrix}.$$

On obtient, partant de la réalisation (ou représentation) $\mathbb{C}$-matricielle de $\mathbb{H}$, une représentation $\mathbb{R}$-matricielle de $\mathbb{H}$. Notons :

$$\mathbf{1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \mathbf{i} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix},$$

$$\mathbf{j} = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \qquad \mathbf{k} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Un élément $h = x\mathbf{1} + y\mathbf{i} + z\mathbf{j} + t\mathbf{k}$ de $\mathbb{H}$, paramétré par quatre réels x, y, z et t, est représenté par :

$$\underline{h} = \left(\begin{array}{cc|cc} x & -y & -z & -t \\ y & x & t & -z \\ \hline z & -t & x & y \\ t & z & -y & x \end{array}\right) = \begin{pmatrix} a & -\overline{b} \\ b & \overline{a} \end{pmatrix},$$

$$\text{où } a = \begin{pmatrix} x & -y \\ y & x \end{pmatrix} \text{ et } b = \begin{pmatrix} z & -t \\ t & z \end{pmatrix}.$$

En d'autres termes, l'application qui, au quaternion $h \in \mathbb{H}$, associe la matrice réelle 4×4 ci-dessus, est une injection de $\mathbb{H}$ dans $\mathscr{M}_4(\mathbb{R})$.

1.4.1. Remarque. Dans l'esprit des remarques 1.2.2 et 1.3.1, la matrice $\underline{h}$ ci-dessus est la matrice de l'application linéaire « multiplication par h », c'est-à-dire $m_h : \mathbb{H} \to \mathbb{H}$, $k \mapsto kh$, dans la base $(\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k})$ de l'espace vectoriel réel $\mathbb{H}$.

2. Interlude : l'ours mal peigné

Voici un interlude qui semble à première vue plus proche de la cosmétologie appliquée que du corps des quaternions. On verra que la question qui suit aboutit à la représentation matricielle (réelle) que l'on vient de rencontrer.

> Peut-on peigner un ours dans un espace de dimension n quelconque sans épi ?

Bien entendu, l'ours doit être vu comme une sphère de dimension $n-1$: il s'agit de trouver un champ de vecteurs tangents à la sphère non singulier. Formellement, on cherche une application continue $X : \mathbf{S}^{n-1} \to \mathbb{R}^n$, $p \mapsto X_p$, de sorte que X_p soit un vecteur non nul et tangent à $\mathbf{S}^{n-1}$ en tout point p.

En dimension (ambiante) 3, l'impossibilité de peigner l'ours a été prouvée par Henri Poincaré à la fin du dix-neuvième siècle. Il est de même impossible de le faire en toute dimension impaire, ce qui a été prouvé par John Milnor en 1978 — c'est le théorème de la boule chevelue. Restait à compter, en dimension paire, le nombre de façons de peigner l'ours, c'est-à-dire le nombre maximal $p(n)$ de champs de vecteurs non singuliers linéairement indépendants. La réponse définitive est donnée par un théorème dû à Adams [1], qui date de 1962. La formule est subtile, mais l'on en déduit l'inégalité : $p(n) \leqslant n-1$, avec égalité pour les seules dimensions 2, 4 et 8.

On se contente ici d'exhiber $p(n)$ champs indépendants pour $2 \leqslant n \leqslant 4$...

Dans $\mathbb{C}$

Topologiquement, la fourrure d'un ours de dimension 2 s'identifie à la sphère-cercle $\mathbf{S}^1$.

On voit bien comment tracer un champ de vecteurs tangents uniforme sur le cercle, l'ours est alors bien peigné : on prend pour chaque vecteur non nul ${}^t(x,y) \in \mathbf{S}^1$ le vecteur tangent ${}^t(-y,x)$, qui ne s'annule jamais.

Y en a-t-il qui nous voient venir ? Si l'on juxtapose un point de la sphère et le choix particulier (section de fibré) d'un vecteur tangent de norme 1, on obtient la matrice

$$\begin{pmatrix} x & -y \\ y & x \end{pmatrix},$$

qui est exactement celle de la représentation réelle des complexes.

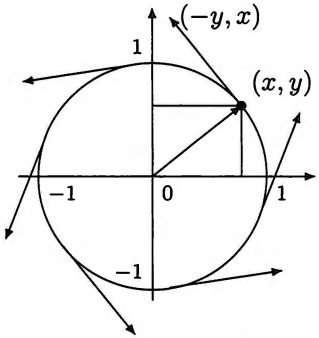


Figure 2.1. *Un ours de dimension* 1 *bien peigné*

Dans $\mathbb{R}^3$

Topologiquement, la fourrure d'un ours de dimension 3 est la sphère $\mathbf{S}^2$.

Il est impossible de lui associer un champ de vecteurs tangents uniformes sans créer un pôle nul, l'ours a alors au moins un épi. D'ailleurs, si c'est une tragédie pour le plantigrade, c'est un bonheur pour le mathématicien, puisque l'on peut montrer grâce à cela que $\mathbb{C}$ est algébriquement clos.

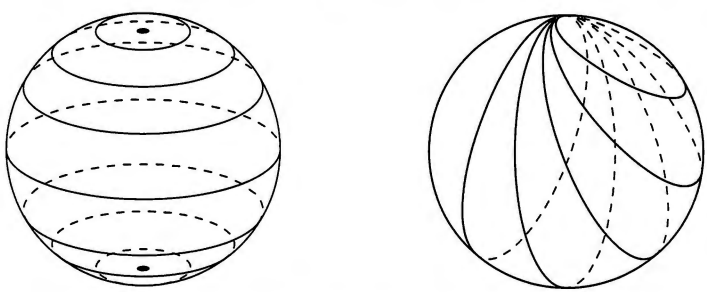


Figure 2.2. *Champs de vecteurs tangents ayant respectivement deux pôles et un pôle*

2.1. Remarque. Une projection stéréographique, voir section X-3.1, centrée sur l'unique pôle de la seconde sphère, permettrait de retrouver toutes les droites du plan d'une direction donnée.

Dans $\mathbb{H}$

Topologiquement, la fourrure d'un ours de dimension 3 est l'hypersphère $\mathbf{S}^3$. Il est possible de lui associer, de $p(4) =$ trois manières différentes, un champ de vecteurs tangents uniforme. L'ours est alors bien peigné : on prend en

chaque vecteur non nul ${}^t(x,y,z,t)$ le vecteur tangent ${}^t(-y,x,-t,z)$, ou ${}^t(-z,t,x,-y)$, ou ${}^t(-t,-z,y,x)$, qui ne s'annulent pas.

À nouveau, si l'on juxtapose le point de la sphère et les trois vecteurs tangents choisis, on retrouve la représentation matricielle réelle des quaternions.

En dimension 8

Le cas $n = 8$ du théorème d'Adams peut être réalisé dans l'espace de dimension 8 des octonions, qui engendre stupeur et tremblements quand on sait que la chose est munie d'une multiplication non commutative et non associative... Voir l'exercice C.5.

3. Applications à SO(3)

C'est ici que l'on comprendra pourquoi $\mathbb{H}$ fut le Graal d'Hamilton : un outil puissant pour faire de la géométrie en dimension 3.

La norme $N(h) = h\overline{h}$ est une forme quadratique réelle définie positive sur $\mathbb{H}$; la base $(\mathbf{1},\mathbf{i},\mathbf{j},\mathbf{k})$ est orthonormée et la forme bilinéaire symétrique associée est donnée par :

$$\forall h,h' \in \mathbb{H}, \quad \langle h,h' \rangle = \frac{1}{2}(h\overline{h'} + h'\overline{h}).$$

(On l'a sans calcul par unicité de la forme bilinéaire symétrique associée.)

Notons que le sous-espace $\mathbb{I}$ des imaginaires de $\mathbb{H}$ est l'orthogonal de $\mathbb{R} = \mathbb{R}\mathbf{1}$ et que $\mathrm{SU}(2) \simeq \mathbf{S}^3$ agit sur $\mathbb{H}$ par automorphismes d'algèbres :

$$\begin{array}{rcl} \varphi : \mathrm{SU}(2) & \longrightarrow & \mathrm{Aut}(\mathbb{H}) \\ h & \longmapsto & \varphi_h : \mathbb{H} \longrightarrow \mathbb{H} \\ & & u \longmapsto huh^{-1}. \end{array}$$

L'application φ_h est linéaire et respecte la norme de $\mathbb{H}$, car $N(huh^{-1}) = N(u)$. Comme $\mathbf{1}$ est central dans $\mathbb{H}$, l'action de $\mathrm{SU}(2)$ préserve $\mathbb{R}$, donc elle préserve aussi son orthogonal $\mathbb{I}$. Considérons alors :

$$\begin{array}{rcl} \varphi : \mathrm{SU}(2) & \longrightarrow & \mathrm{O}(\mathbb{I}) \\ h & \longmapsto & \varphi_h : \mathbb{I} \longrightarrow \mathbb{I} \\ & & u \longmapsto huh^{-1}. \end{array}$$

Via le choix d'une base, qui donne un isomorphisme entre les isométries de $\mathbb{I}$ et le groupe orthogonal $\mathrm{O}(3)$, on définit un morphisme, noté avec un abus par la même lettre : $\varphi : \mathrm{SU}(2) \to \mathrm{O}(3)$.

Par continuité (on n'utilise que des additions, des multiplications et des inverses non nuls), on peut affirmer que l'image $\varphi(\mathrm{SU}(2))$ est connexe et, comme elle contient l'identité, que l'on dispose d'un morphisme φ qui applique SU(2) dans SO(3).

Montrons que ce morphisme est surjectif.

3.1. Rappel. Le résultat suivant est prouvé dans l'annexe A.

- Un système de générateurs de $\mathrm{O}(n)$ est donné par les réflexions orthogonales, qui ont pour matrice, dans une base orthonormée convenable :

$$\begin{pmatrix} \mathrm{I}_{n-1} & \\ & -1 \end{pmatrix}.$$

- Un système de générateurs de $\mathrm{SO}(n)$ est donné par les retournements (demi-tours), qui ont pour matrice, dans une base orthonormée convenable :

$$\begin{pmatrix} \mathrm{I}_{n-2} & \\ & -\mathrm{I}_2 \end{pmatrix}.$$

Pour la surjectivité, il suffit donc de montrer que tous les retournements de $\mathrm{SO}(\mathbb{I}) \simeq \mathrm{SO}(3)$ sont dans $\varphi(\mathrm{SU}(2))$.

Soit $h \in \mathbb{I} \cap \mathbf{S}^3$ et r_h le retournement d'axe $\mathbb{R}h$. On montre que $r_h = \varphi_h$. Pour cela, il suffit de montrer que l'on a :

1. $\varphi_h(h) = h$,
2. $\varphi_h(h') = -h'$ si $\langle h, h' \rangle = 0$.

La première assertion est claire, car l'on a : $\varphi_h(h) = hhh^{-1} = h$.

Quant à la deuxième, on fixe h' orthogonal à h. Il vient : $h'\overline{h} + h\overline{h'} = 0$ donc, puisque h est imaginaire pur : $h'(-h) + h(-h') = 0$, et de là, facilement : $hh'h^{-1} = -h'$.

Donc, l'application $\varphi : \mathrm{SU}(2) \to \mathrm{SO}(3)$ est surjective.

De plus, le noyau de φ est l'intersection du centre de $\mathbb{H}$ avec la sphère unité, il est donc isomorphe, par 1.1.8, au groupe des réels de norme 1. Il vient donc $\mathrm{Ker}\,\varphi = \{\pm \mathrm{I}_2\}$, d'où l'isomorphisme (par passage au quotient) apparaissant dans la proposition suivante.

3.2. Proposition. *Il existe un* isomorphisme exceptionnel *explicite de groupes*

$$\overline{\varphi} : \mathrm{SU}(2)/\{\pm \mathrm{I}_2\} \simeq \mathrm{SO}(3).$$

De plus :

$$\overline{\varphi} : \mathbb{H}^*/\mathbb{R}^* \simeq S^3 \simeq \mathrm{SU}(2).$$

SLU PRODUIT DE L'ANNÉE

3.3. Remarque

- On a en prime une interprétation topologique : SO(3) est homéomorphe à $\mathbb{P}^3(\mathbb{R})$, quotient $\mathbf{S}^3/\{\pm \mathrm{I}_3\}$ de la sphère $\mathbf{S}^3$ par l'antipode.
- Application en calcul formel, et dans les logiciels de simulation de vol, où les calculs de rotations sont effectués dans $\mathbb{H}$ plutôt que dans SO(3) directement.
- Voici un supplément pour les lecteurs qui ont quelques bases de topologie algébrique. Comme SU(2) est la sphère $\mathbf{S}^3$, donc simplement connexe, et comme $\{\mathrm{I}_2, -\mathrm{I}_2\}$ est discret, cela prouve que SU(2) est le revêtement universel de SO(3) et donc que le groupe fondamental de SO(3) est $\mathbb{Z}/2\mathbb{Z}$. En illustration, on a le fameux principe de l'assiette à soupe : si l'on pose sur sa main une assiette et que l'on fait faire à son bras deux rotations de 360^o dans le même sens en maintenant l'assiette horizontale, le bras ne se tord pas deux fois, mais revient à sa position initiale. En aucun cas, la rédaction n'est responsable d'accidents survenus en expérimentant ce beau théorème de topologie.

4. Applications à SO(4)

Avec un tout petit effort supplémentaire, nous allons aussi réaliser SO(4). Bien sûr, cette fois-ci, SU(2) tout seul, trop petit, ne va pas suffire. On considère donc l'action de $\mathrm{SU}(2) \times \mathrm{SU}(2)$ sur $\mathbb{H}$ par automorphismes :

$$\begin{array}{rccl} \psi : \mathrm{SU}(2) \times \mathrm{SU}(2) & \longrightarrow & \mathrm{Aut}(\mathbb{H}) & \\ (h,k) & \longmapsto & \psi_{h,k} : \mathbb{H} \longrightarrow \mathbb{H} & \\ & & u \longmapsto huk^{-1}. & \end{array}$$

Comme avant, on obtient un morphisme : $\psi : \big(SU(2) \times \mathrm{SU}(2)\big) \to \mathrm{SO}(4)$. Montrons qu'il est surjectif.

Soit P un plan quelconque de $\mathbb{H}$, on veut montrer que le retournement par rapport à P est dans l'image de ψ. Pour cela, choisissons une base orthonormée (p,q) de P. On calcule :

$$\overline{p^{-1}q} = \overline{q}\,\overline{p^{-1}} = \overline{q}\,p = -\overline{p}\,q = -p^{-1}q.$$

Par suite, $v := p^{-1}q \in \mathbb{I}$.

D'après le paragraphe sur SO(3), cela implique que $\psi_{v,v}$ est un retournement, et donc que son conjugué $\psi_{p,1}\psi_{v,v}\psi_{p^{-1},1} = \psi_{pvp^{-1},v}$ est aussi un retournement.

De plus, on vérifie facilement que $\psi_{pvp^{-1},v}(p) = p$ et $\psi_{pvp^{-1},v}(q) = q$. C'est donc bien le retournement cherché.

Enfin, on vérifie facilement que le noyau de ψ est : $\{(I_2, I_2), (-I_2, -I_2)\}$.
On a ainsi démontré la proposition suivante.

4.1. Proposition. *Il existe un* isomorphisme exceptionnel *explicite :*

$$\mathrm{SU}(2) \times \mathrm{SU}(2)/\{(I_2, I_2), (-I_2, -I_2)\} \simeq \mathrm{SO}(4).$$

On en déduit des isomorphismes :

$$\begin{aligned}\mathrm{PSO}(4) &\simeq \mathrm{SU}(2) \times \mathrm{SU}(2)/\{(I_2, I_2), (I_2, -I_2), (-I_2, I_2), (-I_2, -I_2)\} \\ &\simeq \mathrm{SU}(2)/\{I_2, -I_2\} \times \mathrm{SU}(2)/\{I_2, -I_2\}.\end{aligned}$$

Et finalement, on obtient l'*isomorphisme exceptionnel* :

$$\mathrm{PSO}(4) \simeq \mathrm{SO}(3) \times \mathrm{SO}(3).$$

Topologiquement, c'est remarquable. Il convient de fêter dignement l'événement en brandissant deux assiettes à soupe, une dans chaque main, en hommage au groupe fondamental de PSO(4).

Algébriquement, c'est une tragédie. Contrairement à tous ses homologues $\mathrm{PSO}(n)$ ($n = 3$ ou $n \geqslant 5$), le groupe PSO(4) n'est pas simple. Il possède des sous-groupes distingués non triviaux, entre autres $\mathrm{SO}(3) \times \{1\}$. C'est d'ailleurs le moment de faire le point sur la simplicité...

5. Aparté sur la simplicité

Mais la beauté du plan réside dans sa simplicité.
The Big Lebowski, Joel & Ethan Coen, 1998.

5.1. Question. Qu'est-ce que la simplicité, au fait ?

Un groupe G est dit simple si ses seuls sous-groupes distingués (stables par conjugaison, *i.e.* par tout automorphisme intérieur) sont le sous-groupe trivial et G lui-même.

5.2. Question. À quoi sert la simplicité, au fait ?

On rappelle qu'un groupe est dit *simple* s'il n'a aucun sous-groupe distingué non trivial. Dans la pratique, on utilise souvent le fait qu'un morphisme non trivial d'un groupe simple vers un autre est injectif. Bien sûr, puisque son noyau est distingué. Si l'on pense par exemple à l'énorme système d'équations que peut représenter le calcul d'un noyau, on comprend rapidement que la simplicité fournit une belle économie de moyens (tout en restant dans l'élégance algébrique).

5.3. Question. Quels sont les groupes simples que l'on est censé connaître au niveau de l'agrégation ?

Réponse :
- le groupe alterné $\mathfrak{A}_n$, *i.e.* le noyau de la signature dans le groupe des permutations $\mathfrak{S}_n$, pour $n \geqslant 5$;
- le quotient $\mathrm{PSL}_n(\mathbb{K})$ du groupe spécial linéaire par son centre pour $(n, \mathbb{K}) \neq (2, \mathbb{F}_2)$ et $(2, \mathbb{F}_3)$;
- le quotient $\mathrm{PSO}(n) = \mathrm{PSO}_n(\mathbb{R})$ du groupe spécial orthogonal par son centre, pour $n \leqslant 3$ ou $n \geqslant 5$.

5.4. Question. Comment montre-t-on qu'un groupe est simple ?

Généralement, on n'en a pas besoin, car les groupes simples sont connus comme le loup blanc...

Mais, la méthode consiste la plupart du temps à montrer qu'un sous-groupe distingué non trivial N du groupe G (dont on veut montrer la simplicité) contiendrait un système de générateurs de G. C'est encore mieux si le système de générateurs forme une (ou deux) classe(s) de conjugaison : il suffit alors de montrer qu'un seul (ou deux) générateur(s) soi(en)t dans N pour que la stabilité par conjugaison implique que tout le système s'y trouve. C'est le cas pour $\mathfrak{A}_n$, engendré par les 3-cycles, $\mathrm{PSL}_n(\mathbb{K})$, engendré par les classes de transvections, et $\mathrm{PSO}(n)$, engendré par les classes de retournements.

Mais, techniquement, c'est souvent redoutable : les gens qui ont inventé la simplicité ne se doutaient pas à quel point cela allait être complexe, alors que ceux qui ont inventé les complexes n'avaient peut-être pas saisi leur simplicité...

Il existe un critère de simplicité fondé sur la table des caractères du groupe (voir le second tome), mais il ne simplifie pas grand-chose.

On signale toutefois un critère dû à Iwasawa qui demande d'exhiber une action du groupe sur un ensemble vérifiant un nombre raisonnable de propriétés.

On pourra se référer à [29] pour la simplicité des groupes cités plus haut.

Références. On pourra aussi se référer à [29], chapitre VII ; [15]. Le morphisme $\mathrm{SU}(2) \to \mathrm{SO}(3)$ est présenté dans [25].

A. Annexe. Engendrement de $O(n)$ et $SO(n)$, simplicité de $SO(3)$

Nous allons donner la preuve de quelques propriétés classiques du groupe $SO(3)$ utilisées dans ce chapitre. Tout d'abord une petite propriété générale et très classique sur l'engendrement des groupes $O(n)$ et $SO(n)$ et sur leurs centres respectifs.

A.1. Proposition. *Le groupe $O(n)$ est engendré par les réflexions orthogonales. Le groupe $SO(n)$ est engendré par les retournements orthogonaux. Leur centre est constitué de leurs homothéties. Comme groupes topologiques, $O(n)$ est compact et $SO(n)$ est compact et connexe.*

Démonstration. La dernière assertion a déjà été prouvée au deuxième chapitre.

Montrons donc les propriétés d'engendrement. On procède par récurrence en commençant par la première dimension sérieuse, la dimension 2. On sait alors qu'un endomorphisme orthogonal est soit une réflexion, soit une rotation et qu'une rotation est la composée de deux réflexions. Cela prouve l'assertion dans le cas où $n = 2$.

Soit n un entier supérieur à 3. Supposons que $O(n-1)$ est engendré par ses réflexions. Soit u dans $O(n)$ distinct de l'identité. Il existe donc x dans l'espace $\mathbb{R}^n$ tel que $u(x) \neq x$. On peut donc définir l'hyperplan médiateur $H := \mathbb{R}\big(x - u(x)\big)^{\perp}$ et s_H, la réflexion par rapport à H. Alors, en remarquant que la décomposition de $u(x)$ sur $H \oplus H^{\perp}$ est donnée par :

$$u(x) = \frac{u(x) - x}{2} + \frac{u(x) + x}{2}$$

(on peut faire un calcul ou, mieux, dessiner un losange !), il vient :

$$\begin{aligned}(s_H \circ u)(x) = s_H\big(u(x)\big) &= s_H\Big(\frac{u(x) - x}{2} + \frac{u(x) + x}{2}\Big) \\ &= -\frac{u(x) - x}{2} + \frac{u(x) + x}{2} = x.\end{aligned}$$

Donc, l'isométrie $s_H \circ u$ fixe x et fixe donc aussi son orthogonal $H' = x^{\perp}$ (H' est un hyperplan, puisque $x \neq 0$). Maintenant, la restriction $\overline{s_H \circ u}$ à cet hyperplan est un élément de $O(n-1)$. Par récurrence, elle est donc produit de réflexions et

$$\overline{s_H \circ u} = \overline{s}_1 \circ \overline{s}_2 \circ \cdots \circ \overline{s}_k.$$

Chaque $\overline{s}_i$ est une réflexion de H', mais elle peut être prolongée en une unique réflexion orthogonale s_i qui fixe x (utiliser le fait que s est une

réflexion orthogonale si et seulement si elle s'écrit $\mathrm{diag}(1,\dots,1,-1)$ dans une base orthonormée). Maintenant, comme $s_H \circ u$ ainsi que $s_1 \circ s_2 \circ \cdots \circ s_k$ coïncident sur H et sur $\mathbb{R}x$, il vient :

$$s_H \circ u = s_1 \circ s_2 \circ \cdots \circ s_k,$$

et donc

$$u = s_H \circ s_1 \circ s_2 \circ \cdots \circ s_k.$$

Cela prouve donc le résultat pour $\mathrm{O}(n)$.

Prouvons maintenant l'assertion pour $\mathrm{SO}(n)$. Soit u dans $\mathrm{SO}(n)$; d'après ce qui précède, u peut s'écrire comme produit de k réflexions. En prenant le déterminant, on voit que k est pair ; de plus, on peut supposer que deux réflexions qui se succèdent dans le produit sont distinctes. Il suffit donc de montrer l'assertion suivante.

Assertion. *Soient H et H' deux hyperplans distincts, alors il existe un couple (r, r') de réflexions tel que $s_H \circ s_{H'} = r \circ r'$.*

Il reste donc à montrer cette assertion. Tout d'abord, posons $F = H \cap H'$. Les hyperplans H et H' étant distincts, on a $H + H' = R^n$ et donc $\dim F = \dim H + \dim H' - \dim(H + H') = n - 2$, par la formule de Grassmann. Soit donc $(e_1, \dots, e_{n-2})$ une base de F, que l'on complète en une base orthonormée $(e_1, \dots, e_n)$ de $\mathbb{R}^n$. On définit alors r et r' par

$$r(e_i) = e_i,\ 1 \leqslant i \leqslant n-3,$$

$$r(e_{n-2}) = -e_{n-2},\ r(e_{n-1}) = -s_H(e_{n-1}),\ r(e_n) = s_H(e_n),$$

$$r'(e_i) = e_i,\ 1 \leqslant i \leqslant n-3,$$

$$r'(e_{n-2}) = -e_{n-2},\ r'(e_{n-1}) = -s_{H'}(e_{n-1}),\ r'(e_n) = s_{H'}(e_n).$$

On a : $s_H \circ s_{H'} = r \circ r'$, car l'égalité est valable sur F ainsi que sur $F^\perp$. Il reste à montrer que r et r' sont des retournements.

Remarquons pour cela que $r \mid F^\perp = s_H \mid F^\perp$, qui est un endomorphisme orthogonal en dimension 2 ayant une unique droite propre de valeur propre -1, qui n'est autre que $H^\perp$: il s'agit donc d'une réflexion orthogonale. En conclusion, r est bien un retournement, et il en est de même pour r'.

Soit z dans le centre de $\mathrm{O}(n)$ et montrons que z est une homothétie (en fait ce sera donc I_n ou $-\mathrm{I}_n$). Soit D une droite quelconque, et soit s la réflexion par rapport à l'hyperplan $D^\perp$. Comme z commute avec s, l'endomorphisme z stabilise tout espace propre de s (résultat classique et facile à montrer) et donc z stabilise D. Comme z stabilise toute droite, z est une homothétie.

De la même manière, si z est dans le centre de SO(n), alors z stabilise tout plan et comme toute droite peut être réalisée comme intersection de deux plans, z stabilise toute droite, si bien que z est une homothétie. □

A.2. Remarques

- L'idée de la preuve en quatre mots : adjonction de points fixes. Le lien avec l'engendrement des groupes symétriques par les transpositions est-il apparent ?
- En étant plus attentif, on peut montrer que tout élément de O(n) est le produit d'au plus n réflexions.
- Le centre de O(n) est donc $\{\mathrm{I}_n, -\mathrm{I}_n\}$. Le centre de SO($n$) est donc $\{\mathrm{I}_n, -\mathrm{I}_n\}$ si n est pair et $\{\mathrm{I}_n\}$ si n est impair.

Suit une preuve élégante de la simplicité de SO(3), fondée sur sa connexité.

A.3. Proposition. *Le groupe* SO(3) *est simple*[7].

Démonstration. On se fixe donc un sous-groupe H non trivial distingué de SO(3) et l'on doit montrer que $H = \mathrm{SO}(3)$. Il suffit de montrer que le sous-groupe H contient un retournement. En effet, si H contient un retournement r_D d'axe D, alors pour toute autre droite D', en notant s un élément de SO(3) envoyant D sur D' (le groupe SO(3) agit transitivement sur la sphère S^2, donc transitivement sur l'ensemble des droites de l'espace). Alors, sr_Ds^{-1} admet le même spectre que r_D et D' est espace propre pour la valeur propre 1, c'est donc le retournement d'axe D' et il est dans H, car H est distingué. Conclusion, si H contient un retournement, il les contient tous, et c'est donc SO(3), puisque celui-ci est engendré par les retournements.

Reste à trouver un retournement dans H. Soit $h \neq \mathrm{I}_3$ dans H. On définit :

$$\varphi : \mathrm{SO}(3) \to \mathbb{R},\ \ \varphi(g) = \mathrm{tr}(ghg^{-1}h^{-1}),$$

où tr désigne la trace. Or, la trace d'un élément de SO(3) est clairement de la forme $1 + 2\cos\theta$. Comme SO(3) est connexe, compact et contient l'identité, son image par φ est de la forme $[a, 3]$, où $a \leqslant 3$.

Supposons par l'absurde $a = 3$, alors cela donnerait $ghg^{-1}h^{-1} = \mathrm{I}_3$ pour tout g de SO(3) et donc h central, c'est-à-dire que h est trivial, contrairement à l'hypothèse.

Donc, $a < 3$. Soit n tel que $a < 1 + 2\cos(\frac{\pi}{n}) < 3$; un tel n existe puisque 3 est la limite de la suite strictement croissante $\big(1 + 2\cos(\pi/n)\big)_{n\in\mathbb{N}^*}$. Soit

[7] Toute preuve utilise à un moment ou un autre le fait que le corps $\mathbb{R}$ est archimédien. Au lecteur de trouver cela dans la démonstration qui suit.

alors g_n dans SO(3) tel que $\varphi(g_n) = 1 + 2\cos(\pi/n)$, ce qui implique que $h_n := g_n h g_n^{-1} h^{-1}$ est une rotation de trace $1 + 2\cos(\pi/n)$, donc d'angle $\pm\pi/n$. De plus, h_n est dans H, car H est distingué. Ainsi, h_n^n est une rotation d'angle π de H, donc un retournement de H comme voulu. □

B. Développement. Fibration de Hopf

B.1. Première version : avec les quaternions

On considère la sphère $\mathbf{S}^2$, (resp. $\mathbf{S}^3$), que l'on réalise comme la sphère unité dans $\mathbb{I}$, (resp. $\mathbb{H}$). Encore une fois, on identifie le groupe SU(2) au groupe des quaternions de norme 1 et à la sphère $\mathbf{S}^3$ de $\mathbb{H}$. Considérons l'action

$$\begin{aligned} \phi : \mathrm{SU}(2) \times \mathbf{S}^2 &\longrightarrow \mathbf{S}^2 \\ (h, z) &\longmapsto \varphi_h(z) = hzh^{-1}. \end{aligned}$$

On voit facilement que SU(2) agit transitivement sur la sphère, puisque son action, qui se factorise par le morphisme surjectif $\varphi : \mathrm{SU}(2) \to \mathrm{SO}(3)$, « est » celle de SO(3) (sur la sphère).

Un calcul simple montre que le stabilisateur de $\mathbf{i} \in \mathbf{S}^2$ est le groupe $\{h = x\mathbf{1} + y\mathbf{i},\ x^2 + y^2 = 1\}$, qui s'identifie au cercle $\mathbf{S}^1$. Par le théorème d'homéomorphisme, l'application orbitale en $\mathbf{i}$ induit un homéomorphisme

$$\mathbf{S}^3/\mathbf{S}^1 \xrightarrow{\simeq} \mathbf{S}^2,$$

très utile quand on veut comprendre la topologie de $\mathbf{S}^3$. La fibration de Hopf est la composée $\pi : \mathbf{S}^3 \to \mathbf{S}^3/\mathbf{S}^1 \to \mathbf{S}^2$.

Pour savoir ce qu'est une fibration, on consultera [25]. Ici, disons simplement que π est une surjection « localement triviale », c'est-à-dire que la restriction de π à l'image réciproque d'un ouvert V de $\mathbf{S}^2$ assez petit s'identifie par un difféomorphisme à la projection naturelle $V \times \mathbf{S}^1 \to V$ (en particulier, toutes les « fibres » $\pi^{-1}(m)$ ($m \in \mathbf{S}^2$) sont des cercles $\mathbf{S}^1$).

B.2. Deuxième version : par projection stéréographique

On peut aussi voir la fibration de Hopf de façon plus géométrique en utilisant la projection stéréographique, voir le paragraphe X-3.1.

On réalise la sphère $\mathbf{S}^3$ dans $\mathbb{C}^2 = \mathbb{R}^4$ comme

$$\mathbf{S}^3 = \{(z_1, z_2) \in \mathbb{C}^2,\ |z_1|^2 + |z_2|^2 = 1\}.$$

On considère alors l'application continue de $\mathbf{S}^3 \to \mathbb{P}^1(\mathbb{C}) = \mathbb{C} \cup \{\infty\}$ qui à un élément (z_1, z_2) non nul associe le quotient z_2/z_1 (ou ∞ si z_1 est nul).

La projection stéréographique du paragraphe X-3.1 donne un homéomorphisme $\mathbb{P}^1(\mathbb{C}) \to \mathbf{S}^2$, ce qui par composition donne une application continue de $\mathbf{S}^3$ sur $\mathbf{S}^2$ dont on vérifie sans sueur que les fibres sont homéomorphes à $\mathbf{S}^1$. En effet, l'image réciproque d'un élément $a \in \mathbb{C}$ est l'intersection $\mathscr{C}_a$ de la sphère unité et de la droite complexe d'équation $z_2 = az_1$, qui est un plan réel : c'est le cercle formé des couples $\big(e^{i\theta}/(1+|a|^2), ae^{i\theta}/(1+|a|^2)\big)$, où θ parcourt $\mathbb{R}$.

B.2.1. Remarque. Un avantage pédagogique certain de cette construction est de pouvoir la transposer, simplifiée, sur le corps des réels. En effet, si l'on remplace $\mathbb{C}$ par $\mathbb{R}$ dans ce qui précède, on obtient une application continue de

$$\mathbf{S}^1 = \{(x_1, x_2) \in \mathbb{R}^2,\ x_1^2 + x_2^2 = 1\}$$

sur $\mathbb{P}^1(\mathbb{R}) = \mathbb{R} \cup \{\infty\}$ (notons qu'il ne s'agit rien d'autre que $(\cos\theta, \sin\theta) \mapsto \tan\theta$), lui même homéomorphe à $\mathbf{S}^1$. La fibre d'un point est alors homéomorphe à $\mathbf{S}^0 = \{1, -1\}$.

On a ainsi la fibration bien connue

$$\mathbf{S}^1/\mathbf{S}^0 \xrightarrow{\simeq} \mathbf{S}^1,$$

qui peut être ainsi réalisée par l'application tangente,... ou par l'application complexe $z \mapsto z^2$.

Peut-être cette remarque permettra-t-elle au lecteur de mieux apprivoiser la fibration de Hopf.

B.3. Toutes *Dimensions* à portée de la main

Pour bien comprendre, la meilleure façon est certainement de regarder les animations des chapitres 7 et 8 de [2]. On peut en effet *dessiner* cette fibration grâce à la projection stéréographique $p : \mathbf{S}^3 \to \mathbb{R}^3 \cup \{\infty\}$ (remarque X-3.1.13, plus loin). Voici deux figures (voir figure B.1 page suivante), pour appâter le lecteur. Dans la première, on fixe un réel $\alpha \in]0, 1[$ et l'on dessine des fibres correspondant à différentes valeurs de a ayant pour module : $|a| = \alpha$. Un point de la réunion des cercles $\mathscr{C}_{\alpha e^{i\varphi}}$ (φ réel) est paramétré par $(\varphi, \theta) \in (\mathbb{R}/2\pi\mathbb{Z})^2$: la réunion est un tore. En fait, c'est même un tore de révolution sur lequel les $\mathscr{C}_{\alpha e^{i\varphi}}$ forment une des deux familles de « cercles de Villarceau ». Dans la seconde figure, sont représentés plusieurs de ces tores, regroupés par nuance de gris[8].

[8] Très à la mode en cette fin d'année 2012.

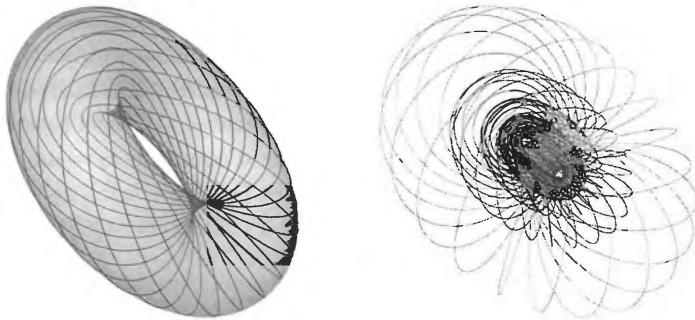


Figure B.1. *Fibration de Hopf*

Références

Pour une introduction à la fibration de Hopf qui n'utilise pas les quaternions, on pourra consulter [25], chapitre 4. Sinon, répétons-le, on pourra aussi contempler les chapitres 7 et 8 du film [2] qui lui sont consacrés.

C. Exercices

Le corps des quaternions et applications

C.1. Exercice ($\mathscr{M}_n(\mathbb{C})$ dans $\mathscr{M}_{2n}(\mathbb{R})$).

Soit n un entier naturel. Définir un morphisme injectif d'algèbres réelles de $\mathscr{M}_n(\mathbb{C})$ dans $\mathscr{M}_{2n}(\mathbb{R})$.

On pourra considérer l'espace vectoriel *complexe* $E = \mathbb{C}^n$ et l'espace vectoriel *réel* $E_{\mathbb{R}} = \mathbb{C}^n$ (voir l'exercice II-F.19) et constater qu'un endomorphisme de E est *a fortiori* un endomorphisme de $E_{\mathbb{R}}$. Observer l'injection $\mathscr{L}(E) \to \mathscr{L}(E_{\mathbb{R}})$ à travers des bases convenables.

C.2. Exercice (Automorphismes de l'algèbre des quaternions)

Le but de l'exercice est de prouver que tout automorphisme de l'algèbre des quaternions est intérieur. Soit donc φ un automorphisme de l'algèbre des quaternions $\mathbb{H}$, alors avec les notations du § 3, on veut montrer qu'il existe h dans $\mathbb{H}\backslash\{0\}$ tel que $\varphi = \varphi_h$.

1. *Montrer que si u est un imaginaire pur,* i.e. *est dans le sous-espace $\mathbb{I}$ engendré par i, j, k, alors $\varphi(u) = -N(u)$.*
2. *Montrer qu'un quaternion est un imaginaire pur si et seulement si son carré est un réel négatif.*

3. *En déduire que* $\mathbb{I}$ *est stable par* φ *et que* $\varphi|_{\mathbb{I}}$ *est dans* $\mathrm{O}(\mathbb{I}) \simeq \mathrm{O}(3)$.

4. *Montrer que si* $\varphi|_{\mathbb{I}}$ *est dans* $\mathrm{SO}(\mathbb{I})$, *alors, on peut trouver un* h *non nul tel que* $\varphi = \varphi_h$.
   Utiliser le § 3.

5. *Montrer que l'endomorphisme de l'espace* $\mathbb{I}$ *tel que* $i \mapsto -i$, $j \mapsto j$, $k \mapsto k$, *n'est pas la restriction d'un endomorphisme de l'algèbre* $\mathbb{H}$. *En déduire que* $\varphi|_{\mathbb{I}}$ *est dans* $\mathrm{SO}(\mathbb{I}) \simeq \mathrm{SO}(3)$ *et conclure.*

C.3. Exercice (Isomorphisme $\mathrm{SU}(2)/\{1,-1\} \simeq \mathrm{SO}_3$, via la projection stéréographique)

Soit $\mathbf{S}^2$ *la sphère unité de* $\mathbb{R}^3$, *d'équation* $x_1^2 + x_2^2 + x_3^2 = 1$. *Soit* φ *la* projection stéréographique $\varphi : \mathbf{S}^2 \setminus \{N\} \longrightarrow \mathbb{C}$, *où* $N = (0,0,1)$. *Elle associe à un point* $x \in \mathbf{S}^2 \setminus \{N\}$ *le point d'affixe* z *du plan horizontal (identifié à* $\mathbb{C}$*), appartenant à la droite joignant* N *et* x.

1. *Montrer que* φ *est donnée par la formule*

$$z = \frac{x_1 + ix_2}{1 - x_3},$$

   et que son inverse φ^{-1} *est donnée par*

$$x_1 + ix_2 = \frac{2z}{|z|^2 + 1}, \qquad x_3 = \frac{|z|^2 - 1}{|z|^2 + 1}.$$

   On fera un dessin et l'on observera une configuration de Thalès.

2. *Si* $g = \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix} \in \mathrm{SU}(2)$, *on note* $T(g)$ *la transformation homographique définie par*

$$T(g)(z) = \frac{\alpha z + \beta}{-\overline{\beta} z + \overline{\alpha}}.$$

   Posons $R(g) := \varphi^{-1} \circ T(g) \circ \varphi$.

   (a) *Expliciter la composée* $\varphi^{-1} \circ T(g) \circ \varphi(x_1, x_2, x_3)$. *En déduire qu'il existe* $f \in \mathscr{G}(\mathbb{R}^3)$ *tel que* $f|_{S_2} = R(g)$

   (b) *En déduire que* R *est un morphisme de* $\mathrm{SU}(2)$ *dans* $\mathrm{SO}(3)$.

3. *Montrer que* $\operatorname{Ker} R = \{\pm \mathrm{I}_2\}$. *En déduire que la différentielle* dR_1 *de* R *en l'identité est un isomorphisme.*

4. *En déduire*[9] *qu'il existe une suite exacte*

$$\{1\} \longrightarrow \mathbb{Z}/2\mathbb{Z} \longrightarrow \mathrm{SU}(2) \longrightarrow \mathrm{SO}(3) \longrightarrow \{1\}.$$

[9] Le lecteur pourra, à cet effet, jeter un regard en page 273 sur la démonstration de la proposition IX-2.1.

C.4. Exercice (Application des quaternions au groupe symplectique). *Soit n un entier non nul. On note J la matrice $2n\times 2n$ donnée par blocs par*

$$J = \begin{pmatrix} 0_n & -\mathrm{I}_n \\ \mathrm{I}_n & 0_n \end{pmatrix}.$$

On note ω la forme bilinéaire antisymétrique de $\mathbb{C}^{2n}$ qui a pour matrice J dans la base canonique. Le groupe $\mathrm{Sp}_{2n}(\mathbb{C})$ est le sous-groupe de $\mathrm{GL}_{2n}(\mathbb{C})$ qui préserve ω. On veut étudier les propriétés de $\mathrm{Sp}_{2n}(\mathbb{C})$ et, en particulier, prouver qu'il est inclus dans $\mathrm{SL}_{2n}(\mathbb{C})$.

1. *Montrer que pour tout A de $\mathrm{Sp}_{2n}(\mathbb{C})$, on a : $\det(A) = \pm 1$.*
   On munit $\mathbb{H}^n$ du produit quaternionique

$$\langle U, V\rangle = \sum_{i=1}^{n} u_i\,\overline{v_i} \in \mathbb{H}, \quad \text{où } U = (u_i), V = (v_i) \in \mathbb{H}^n.$$

   On note $\mathbb{C}$ le sous-corps $\mathbb{R}\oplus\mathbb{R}i$ de $\mathbb{H}$. On remarque que tout élément $u = x + yi + zj + tk$ de $\mathbb{H}$, avec x, y, z et t réels, peut s'écrire de façon unique sous la forme $u = u'_1 + u'_2 j$ avec u'_1 et u'_2 complexes : il suffit de poser $u'_1 = x + yi$ et $u'_2 = z + ti$. On identifiera $\mathbb{H}^n$ à $\mathbb{C}^{2n}$ via l'isomorphisme de décomposition complexe $U = (u_1, \ldots, u_n) \mapsto U' = (u'_1, \ldots, u'_n, u'_{n+1}, \ldots, u'_{2n})$ avec, pour tout ℓ, $u'_\ell \in \mathbb{C}$, $u'_{n+\ell} \in \mathbb{C}$, $u_\ell \in \mathbb{H}$ et $u_\ell = u'_\ell + u'_{n+\ell} j$.

2. *Montrer que pour tout a dans $\mathbb{C}$, on a : $aj = j\overline{a}$. Montrer par un calcul que si V, W sont dans $\mathbb{H}^n$ et si V', W' sont leur image dans $\mathbb{C}^{2n}$, alors*

$$\langle V, W\rangle = (V', W') + \omega(V', W')j,$$

   où $(\cdot\,,\cdot)$ désigne le produit scalaire hermitien usuel de $\mathbb{C}^{2n}$.

3. *On note $\mathrm{Sp}(n) = \mathrm{U}_{2n} \cap \mathrm{Sp}_{2n}(\mathbb{C})$, où U_{2n} désigne le sous-groupe des matrices unitaires et $\mathrm{O}(n, \mathbb{H})$ le groupe des endomorphismes de $\mathbb{H}^n$ (comme $\mathbb{C}$-espace) qui conservent le produit quaternionique. Déduire l'isomorphisme $\mathrm{Sp}(n) \simeq \mathrm{O}(n, \mathbb{H})$.*

4. *Déduire de la question 3 que $\mathrm{Sp}(n)$ agit sur la sphère (réelle) $\mathbf{S}^{4n-1}$ de dimension $4n-1$ et que l'on a un homéomorphisme*

$$\mathrm{Sp}(n)/\mathrm{Sp}(n-1) \simeq \mathbf{S}^{4n-1}.$$

   On pourra généraliser au corps-non-commutatif $\mathbb{H}$ des résultats de bases incomplètes bien connus sur des corps commutatifs. Par exemple, on peut penser à généraliser le procédé de Gram-Schmidt.

5. *Montrer que $\mathrm{Sp}(n)$ est un groupe connexe.*

6.(a) *Soit M une matrice hermitienne définie positive de $\mathrm{Sp}_{2n}(\mathbb{C})$ et H une matrice hermitienne telle que $\exp H = M$. Montrer que l'on a : $HJ + J\,{}^tH = 0$.*

(b) Montrer que si $A \in \mathrm{Sp}_{2n}(\mathbb{C})$, alors : $A^* \in \mathrm{Sp}_{2n}(\mathbb{C})$.

(c) Déduire, à l'aide de la décomposition polaire, l'homéomorphisme

$$\mathrm{Sp}_{2n}(\mathbb{C}) \simeq \mathrm{Sp}(n) \times \mathbb{R}^d,$$

où d est un entier à déterminer.

7. *Montrer que $\mathrm{Sp}_{2n}(\mathbb{C})$ est connexe. Avec ce qui précède, en déduire qu'il est inclus dans $\mathrm{SL}(2n, \mathbb{C})$.*

Instructif

C.5. Exercice (Octonions. Construction de Cayley-Dickson)

Soit A une algèbre sur un corps $\mathbb{K}$ munie d'une involution $A \to A$, $a \mapsto a^*$, telle que :

$$\forall a, b \in A, \quad (ab)^* = b^* a^*.$$

On munit l'espace $A^{[2]} = A \times A$ du produit suivant :

$$\forall (a,b),(c,d) \in A \times A, \quad (a,b) \cdot (c,d) = (ac - d^* b, da + bc^*).$$

1. *On définit le conjugué d'un élément $(a,b) \in A^{[2]}$ par : $(a,b)^* = (a^*, -b)$. Montrer que la conjugaison est un anti-morphisme d'algèbres. Calculer $(a,b)(a,b)^*$.*
2. *Identifier $A^{[2]}$ lorsque :*
   (a) $A = \mathbb{R}$ et $a^* = a$;
   (b) $A = \mathbb{C}$ et $a^* = \overline{a}$.
3. *On appelle algèbre des octonions et l'on note $\mathbb{O}$ l'algèbre $A^{[2]}$ lorsque $A = \mathbb{H}$ et l'involution * est la conjugaison quaternionique. Montrer que tout élément non nul possède un inverse. Montrer que $\mathbb{O}$ est une algèbre qui n'est pas associative, mais qui est alternative, au sens où :*

$$\forall a, b \in \mathbb{O}, \quad a(ab) = (aa)b \text{ et } (ba)a = b(aa).$$

C.6. Exercice (« Bicomplexes » – extension des scalaires pour les complexes)

On a vu au paragraphe 1.2 *comment réaliser $\mathbb{C}$ comme une algèbre réelle C de dimension 2. On peut étendre les scalaires à $\mathbb{C}$, c'est-à-dire considérer le sous-espace vectoriel* complexe *$C_{\mathbb{C}}$ de $\mathscr{M}_2(\mathbb{C})$ engendré par les matrices*

$$\mathbf{1} = \mathbb{I}_2 \quad \text{et} \quad \mathbf{i} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

1. *Vérifier que* $C_{\mathbb{C}}$ *est une sous-algèbre de* $\mathscr{M}_2(\mathbb{C})$.
2. *Montrer que* $C_{\mathbb{C}}$ *contient deux idempotents orthogonaux non triviaux, c'est-à-dire deux éléments* e *et* e' *tels que* $e^2 = e$, $e'^2 = e'$, $ee' = e'e = 0$ *et* $e + e' = \mathbb{I}_2$.

   L'équation matricielle $e^2 = e$ admet essentiellement deux solutions non triviales dans $C_{\mathbb{C}}$; le reste suit.
3. *Montrer que* $C_{\mathbb{C}}$ *est isomorphe à* $\mathbb{C} \oplus \mathbb{C}$.

   L'application $C_{\mathbb{C}} \to \mathbb{C} \oplus \mathbb{C}$, $c \mapsto (ece, e'ce')$ est un isomorphisme.
4. *Exhiber un isomorphisme d'algèbres réelles* $\mathbb{R}[X]/(X^2+1) \to C$ *et montrer qu'il se prolonge en un isomorphisme d'algèbres complexes* $\mathbb{C}[X]/(X^2+1) \to C_{\mathbb{C}}$. *En déduire une deuxième preuve du fait que* $C_{\mathbb{C}}$ *est isomorphe à* $\mathbb{C} \oplus \mathbb{C}$.
5. *Montrer que le* C*-module* C *est simple, c'est-à-dire qu'il n'y a pas de sous-espace vectoriel réel dans* C *stable par multiplication par* C. *En revanche, vérifier que le* $C_{\mathbb{C}}$*-module* $C_{\mathbb{C}}$ *est somme directe de deux sous-modules simples.*

   N.B. On peut comprendre ce phénomène de scindage en termes de représentations de $\mathbb{Z}/4\mathbb{Z}$ *: voir l'exercice* 3.

C.7. Exercice (« Biquaternions » – extension des scalaires pour les quaternions).

On a vu au paragraphe 1.4 *comment réaliser* $\mathbb{H}$ *comme une algèbre réelle* A *de dimension* 4. *On peut étendre les scalaires à* $\mathbb{C}$, *c'est-à-dire considérer le sous-espace vectoriel* complexe $A_{\mathbb{C}}$ *de* $\mathscr{M}_4(\mathbb{C})$ *engendré par les matrices*

$$\mathbf{1} = \mathrm{I}_4,\ \mathbf{i} = \ldots,\ \mathbf{j} = \ldots,\ \mathbf{k} = \ldots.$$

Par ailleurs, on note $\mathbb{H}_{\mathbb{C}}$ *l'espace vectoriel de dimension* 4 *sur* $\mathbb{C}$, *ayant une base* $(\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k})$, *muni de la multiplication donnée par les relations de la définition* 1.1.1.

Pour des raisons évidentes, on dit que $A_{\mathbb{C}}$ *et* $\mathbb{H}_{\mathbb{C}}$ *sont obtenues de* A *et* $\mathbb{H}$ *par extension des scalaires.*

1. *Vérifier que* $\mathbb{H}_{\mathbb{C}}$ *et* $A_{\mathbb{C}}$ *sont des sous-algèbres de* $\mathscr{M}_4(\mathbb{C})$ *et qu'elles sont isomorphes.*
2. *Montrer que l'application suivante, qui étend à la situation complexe l'isomorphisme réel* $\mathbb{H} \to H$ *du paragraphe* 1.4, *est un isomorphisme de* $\mathbb{H}_{\mathbb{C}}$ *sur* $\mathscr{M}_2(\mathbb{C})$ *:*

$$\mathbb{H}_{\mathbb{C}} \to \mathscr{M}_2(\mathbb{C}), \quad x\mathbf{1} + y\mathbf{i} + z\mathbf{j} + w\mathbf{k} \mapsto \begin{pmatrix} w+iy & z+it \\ -z+it & x-iy \end{pmatrix}.$$

3. *Montrer que le A-module A est simple (ses seuls sous-modules sont $\{0\}$ et A), alors que le $A_{\mathbb{C}}$-module $A_{\mathbb{C}}$ est somme directe de deux sous-modules de dimension complexe 2.*

N.B. Ainsi, l'algèbre A n'est pas isomorphe à une algèbre de matrices, alors qu'en étendant les scalaires à $\mathbb{C}$ — c'est-à-dire en considérant $A_{\mathbb{C}}$ — on obtient une algèbre de matrices.

N.B. On verra le même phénomène de scindage qu'à la question 3 à l'œuvre dans l'étude des représentations du « groupe quaternionique »

$$H_8 = \{\pm\mathbf{1}, \pm\mathbf{i}, \pm\mathbf{j}, \pm\mathbf{k}\}$$

(Voir second tome, exercice X-D.9*).*

Chapitre VIII

Combinatoire algébrique

"Rhââ lovely !"

Frenzy, Alfred Hitchcock, 1972.

Les corps finis sont un passage obligé dans grand nombre de problèmes arithmétiques, mais nous n'aborderons malheureusement pas ce point de vue. Dans ce chapitre, nous apportons un nouvel éclairage sur les groupes classiques et la géométrie en les considérant sur les corps finis. On retrouve l'équivalent des groupes déjà étudiés sur $\mathbb{R}$ et $\mathbb{C}$ (groupe linéaire, groupe spécial linéaire, groupe orthogonal), mais cette fois-ci la combinatoire se substitue à la topologie. Les groupes agiront sur des objets géométriques (on parle de « géométrie discrète »), mais la surjectivité ou parfois l'injectivité, pourra être montrée à l'aide du dénombrement.

Le dénombrement des objets du cours possède un premier avantage ; il permet à l'étudiant de mieux comprendre, voire de tester sa propre compréhension, selon le principe que « compter, c'est comprendre ». Par exemple, il n'y a rien de mieux, pour sentir que l'on a dominé la notion de repère projectif, que de compter le nombre de repères projectifs sur les corps finis, et, ainsi, réaliser l'aide qu'apporte une action du groupe adéquat dans ce dénombrement. Le second point positif, c'est que cette approche des objets géométriques est particulièrement jubilatoire : compter, c'est jouer avec les nombres, c'est la magie de l'enfance devant les chiffres, et ici, jointe à une géométrie des corps finis, hautement hallucinogène.

Nous profiterons de cet enthousiasme général pour affiner quelques notions de géométrie projective, qui seront d'ailleurs reprises dans les prochains

chapitres. Nous reverrons donc l'espace projectif, le groupe projectif, et nous introduirons les repères projectifs. Enfin, nous présenterons quelques isomorphismes, dits exceptionnels, entre groupes de nature différente, et pour n'en citer qu'un, le groupe projectif $\mathrm{PSL}_2(\mathbb{F}_5)$, défini comme étant le quotient du groupe $\mathrm{SL}_2(\mathbb{F}_5)$ par ses homothéties est isomorphe au groupe alterné $\mathfrak{A}_5$. Encore une fois, le dénombrement sera essentiel dans l'étude de ces isomorphismes. Les isomorphismes exceptionnels sur $\mathbb{R}$ et $\mathbb{C}$, plus difficiles, seront étudiés dans un chapitre ultérieur, à l'aide non seulement de la topologie, mais d'une structure différentielle.

1. Dénombrement sur les corps finis

Nous allons dénombrer quelques objets courants de la géométrie sur un corps fini $\mathbb{F}_q$.

On rappelle que $\mathbb{P}^{n-1}(\mathbb{F}_q)$ représente l'ensemble des droites (vectorielles) de l'espace $\mathbb{F}_q^n$. Le groupe $\mathrm{PGL}_n(\mathbb{F}_q)$ est le quotient de $\mathrm{GL}_n(\mathbb{F}_q)$ par le groupe $\mathbb{F}_q^*$ de ses homothéties. Sa raison d'être étant d'agir de façon fidèle sur l'espace projectif $\mathbb{P}^{n-1}(\mathbb{F}_q)$. Un repère projectif de l'espace $\mathbb{P}^{n-1}(\mathbb{F}_q)$ est un $(n+1)$-uplet de droites de $\mathbb{F}_q^n$ telles qu'aucune sous-famille à n éléments ne soit dans un même hyperplan de $\mathbb{F}_q^n$. Pour se familiariser avec la notion de repère projectif, il est recommandé de regarder le cas de la dimension 2 (voir notamment en page 333).

1.1. Proposition. *Sur un corps fini $\mathbb{F}_q$ à q éléments et pour un entier n donné, les ensembles suivants ont pour cardinaux :*

(i) *Espace vectoriel :*

$$|\mathbb{F}_q^n| = q^n.$$

(ii) *Espace projectif :*

$$|\mathbb{P}^n(\mathbb{F}_q)| = 1 + q + q^2 + \cdots + q^n.$$

(iii) *Groupe général linéaire :*

$$|\mathrm{GL}_n(\mathbb{F}_q)| = (q^n - 1)(q^{n-1} - 1) \times \cdots \times (q - 1)q^{n(n-1)/2}.$$

C'est aussi le nombre de bases de $\mathbb{F}_q^n$.

(iv) *Groupe projectif linéaire :*

$$|\mathrm{PGL}_n(\mathbb{F}_q)| = (q^n - 1)(q^{n-1} - 1) \times \cdots \times (q^2 - 1)q^{n(n-1)/2}.$$

C'est aussi le nombre de repères de $\mathbb{P}^{n-1}(\mathbb{F}_q)$.

(v) *Groupe spécial linéaire*

$$|\operatorname{SL}_n(\mathbb{F}_q)| = (q^n-1)(q^{n-1}-1)\times\cdots\times(q^2-1)q^{n(n-1)/2}.$$

(vi) *Grassmannienne :*

$$|\operatorname{Gr}_{m,n}(\mathbb{F}_q)| = \frac{(q^n-1)(q^{n-1}-1)\times\cdots\times(q^{n-m+1}-1)}{(q^m-1)(q^{m-1}-1)\times\cdots\times(q-1)}.$$

(vii) *« Variété » de drapeaux :*

$$\mathscr{F}_n(\mathbb{K}) = \{(F_0,\dots,F_n) \,:\, \dim F_i = i,\ \{0\}=F_0\subset\cdots\subset F_n=\mathbb{K}^n\}$$
$$|\mathscr{F}_n(\mathbb{F}_q)| = \prod_{k=1}^{n-1}(1+q+\cdots+q^k).$$

Démonstration. (i) Clair.

(ii) Pour $\mathbb{K}$ corps quelconque, on fait agir le groupe multiplicatif $\mathbb{K}^*$ sur $\mathbb{K}^{n+1}\setminus\{0\}$ par homothéties, c'est-à-dire qu'un scalaire $\lambda\in\mathbb{K}^*$ agit sur un vecteur non nul v par : $\lambda\cdot v = \lambda v$. L'action est libre : pour v vecteur non nul, on a : $\operatorname{Stab}_{\mathbb{K}^*}(v)=\{1\}$. Par suite, pour $\mathbb{K}=\mathbb{F}_q$, on a :

$$|\mathbb{P}^n(\mathbb{F}_q)| = \left|\mathbb{F}_q^{n+1}\setminus\{0\}/\mathbb{F}_q^*\right| = \frac{|\mathbb{F}_q^{n+1}\setminus\{0\}|}{|\mathbb{F}_q^*|}$$
$$= \frac{q^{n+1}-1}{q-1} = 1+q+q^2+\cdots+q^n.$$

Reformulons : les orbites sont les droites de $\mathbb{F}_q^{n+1}$; il y a $q^{n+1}-1$ vecteurs non nuls et une même droite peut être engendrée par $q-1$ vecteurs différents, et les droites sont donc bien au nombre de $(q^{n+1}-1)/(q-1)$.

(iii) Pour $\mathbb{K}$ corps quelconque, on fait agir $\operatorname{GL}_n(\mathbb{K})$ sur $\mathbb{K}^n\setminus\{0\}$. L'action est transitive et le stabilisateur du premier vecteur de la base standard est :

$$\operatorname{Stab}(e_1) = \begin{pmatrix} 1 & \mathbb{K}^{n-1} \\ 0 & \operatorname{GL}_{n-1}(\mathbb{K}) \end{pmatrix} \simeq \operatorname{GL}_{n-1}(\mathbb{K}) \ltimes \mathbb{K}^{n-1}.$$

Pour $\mathbb{K}=\mathbb{F}_q$, on a $\left|\operatorname{GL}_n(\mathbb{F}_q)/\left(\operatorname{GL}_{n-1}(\mathbb{F}_q)\ltimes\mathbb{F}_q^{n-1}\right)\right| = |\mathbb{F}_q^n\setminus\{0\}|$, ce qui donne : $|\operatorname{GL}_n(\mathbb{F}_q)| = |\operatorname{GL}_{n-1}(\mathbb{F}_q)|\cdot q^{n-1}(q^n-1)$.

Par récurrence, on trouve : $|\operatorname{GL}_n(\mathbb{F}_q)| = (q^n-1)(q^n-q)\times\cdots\times(q^n-q^{n-1})$.

Une autre méthode consiste à compter le nombre de bases de $\mathbb{F}_q^n$. En effet, une fois fixée une base de référence, par exemple la base canonique $\mathbf{e} = (e_1,\dots,e_n)$, l'application qui, à un élément $g\in\operatorname{GL}_n(\mathbb{K})$, associe la base $(ge_1,\dots,ge_n)$ est une bijection du groupe linéaire vers l'ensemble des bases.

Son application réciproque est l'application qui à une base $\mathbf{e}'$ de $\mathbb{K}^n$ associe la matrice de passage de $\mathbf{e}$ à $\mathbf{e}'$. En d'autres termes, $\mathrm{GL}_n(\mathbb{K})$ agit sur l'ensemble des bases de façon simplement transitive.

Pour $\mathbb{K} = \mathbb{F}_q$, une base étant une partie libre à n éléments de $\mathbb{F}_q^n$, son premier vecteur est non nul : il y a donc $q^n - 1$ possibilités.

Une fois choisi le premier, le second vecteur doit être en-dehors de la droite qu'il engendre, il y a $q^n - q$ possibilités. On continue jusqu'au n^{e} vecteur, non-colinéaire aux $n-1$ premiers vecteurs et donc au choix parmi $q^n - q^{n-1}$ vecteurs.

Au bilan, on obtient : $|\mathrm{GL}_n(\mathbb{F}_q)| = (q^n - 1)(q^n - q)\times\cdots\times(q^n - q^{n-1})$.

(iv) On a la suite exacte :

$$1 \longrightarrow \mathbb{K}^* \longrightarrow \mathrm{GL}_n(\mathbb{K}) \longrightarrow \mathrm{PGL}_n(\mathbb{K}) \longrightarrow 1.$$

D'où : $\mathrm{PGL}_n(\mathbb{K}) \simeq \mathrm{GL}_n(\mathbb{K})/\mathbb{K}^*$ et, pour $\mathbb{K} = \mathbb{F}_q$:

$$|\mathrm{PGL}_n(\mathbb{F}_q)| = \frac{|\mathrm{GL}_n(\mathbb{F}_q)|}{|\mathbb{F}_q^*|} = (q^n - 1)(q^{n-1} - 1)\times\cdots\times(q^2 - 1)q^{n(n-1)/2}.$$

On peut aussi compter le nombre de repères projectifs de $\mathbb{P}^{n-1}$ (voir la proposition XI-2.3.3), puisque le groupe agit de façon simplement transitive sur ceux-ci. Un calcul du même genre que celui de (iii) donne :

$$|\mathrm{PGL}_n(\mathbb{F}_q)| = \underbrace{(1 + q + \cdots + q^{n-1})}_{\in\mathbb{P}^{n-1}}\underbrace{(q + \cdots + q^{n-1})}_{\in\mathbb{P}^{n-1}\setminus\{1\}}\times\cdots\times\underbrace{q^{n-1}}_{\in\mathbb{P}^{n-1}\setminus\mathbb{P}^{n-2}}(q-1)^{n-1}.$$

(v) De même, pour tout corps, on a la suite exacte

$$1 \longrightarrow \mathrm{SL}_n(\mathbb{K}) \longrightarrow \mathrm{GL}_n(\mathbb{K}) \longrightarrow \mathbb{K}^* \longrightarrow 1.$$

D'où :

$$\mathrm{Im}(\det) = \mathbb{K}^* \simeq \mathrm{GL}_n(\mathbb{K})/\operatorname{Ker}(\det) = \mathrm{GL}_n(\mathbb{K})/\mathrm{SL}_n(\mathbb{K})$$

et, pour $\mathbb{K} = \mathbb{F}_q$: $|\mathrm{SL}_n(\mathbb{F}_q)| = |\mathrm{GL}_n(\mathbb{F}_q)|/|\mathbb{F}_q^*|$.

(vi) Comme cela a été vu, $\mathrm{GL}_n(\mathbb{K})$ agit transitivement sur la grassmanienne $\mathrm{Gr}_{m,n}(\mathbb{K})$, voir proposition II-4.3.9. Le stabilisateur d'un sous-espace est donné par :

$$\begin{aligned}\mathrm{Stab}(F) &= \{g \in \mathrm{GL}_n(\mathbb{K}) : g(f_i) = \lambda(f_i),\ \lambda \in \mathbb{K}\}\\ &= \begin{pmatrix} \mathrm{GL}_m(\mathbb{K}) & \mathscr{M}_{m,n-m}(\mathbb{K}) \\ 0 & \mathrm{GL}_{n-m}(\mathbb{K}) \end{pmatrix}\\ &= \big(\mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_{n-m}(\mathbb{K})\big) \ltimes \mathscr{M}_{m,n-m}(\mathbb{K}).\end{aligned}$$

Ainsi, pour $\mathbb{K} = \mathbb{F}_q$, on obtient :

$$|\operatorname{Gr}_{m,n}(\mathbb{F}_q)| = \frac{|\operatorname{GL}_n(\mathbb{F}_q)|}{|\operatorname{GL}_m(\mathbb{F}_q)| \times |\operatorname{GL}_{n-m}(\mathbb{F}_q)| \times |\mathscr{M}_{m,n-m}(\mathbb{F}_q)|}.$$

(vii) Par le théorème de la base incomplète (encore lui !), $\operatorname{GL}_n(\mathbb{K})$ agit transitivement sur les variétés de drapeaux $\mathscr{F}_n(\mathbb{K})$, et le stabilisateur du drapeau standard est le groupe des matrices triangulaires supérieures inversibles.

Détaillons un peu. Le drapeau standard est le drapeau dont le k^{e} sous-espace F_k^0 est engendré par les k premiers vecteurs de la base canonique. Étant donné un drapeau $(F_k)_k$, on construit de proche en proche une base qui lui est compatible, c'est-à-dire telle que $(e_1, \cdots, e_k)$ est une base de F_k. La matrice de passage g de la base canonique à cette base envoie F_k^0 sur F_k pour tout k. Stabiliser le drapeau standard, c'est envoyer chaque e_k dans F_k, c'est-à-dire sur une combinaison linéaire des e_ℓ, $\ell \leqslant k$.

Ainsi, le stabilisateur est en bijection[1] avec $(\mathbb{K}^*)^n \times \mathbb{K}^{n(n-1)/2}$ et son cardinal, pour $\mathbb{K} = \mathbb{F}_q$, est : $(q-1)^n q^{n(n-1)/2}$. Donc :

$$|\mathscr{F}_n(\mathbb{F}_q)| = \frac{|\operatorname{GL}_n(\mathbb{F}_q)|}{(q-1)^n q^{n(n-1)/2}}.$$

□

1.2. Remarque. On montrera mieux au second volume, dans § IV-1 : l'expression de $|\operatorname{Gr}_{m,n}(\mathbb{F}_q)|$ est un polynôme $P_{m,n}(q)$ tel que $P_{m,n}(1) = \binom{n}{m}$. Ce polynôme est un « équivalent polynomial », ou « analogue quantique », des nombres binomiaux.

Voici quelques exemples.

1.3. Exemple. Le nombre d'éléments de $\mathbb{P}^1(\mathbb{F}_2)$ peut se déterminer graphiquement en regardant les droites de $\mathbb{F}_2^2$.

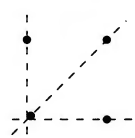


Figure 1.1. *Points de* $\mathbb{F}_2^2 \setminus \{0\}$ *et droites de* $\mathbb{P}^1(\mathbb{F}_2)$

Le nombre de droites éléments de $\mathbb{P}^1(\mathbb{F}_3)$ peut se déterminer graphiquement à partir des droites de $\mathbb{F}_3^2$ (figure 1.2).

[1] Attention, les groupes ne sont pas isomorphes, même en remplaçant $\times$ par $\ltimes$.

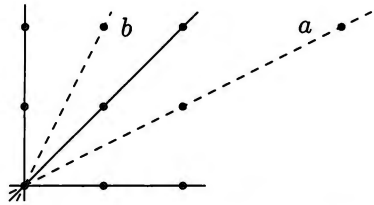


Figure 1.2. *Points de $\mathbb{F}_3^2$ et droites de $\mathbb{P}^1(\mathbb{F}_3)$*
(Les droites pointillées forment en fait une seule,
car $a = b$ — et que par deux points passe une droite et une seule !)

À titre d'exemple, nous avons représenté les éléments de $\mathbb{F}_p^2$ dans le plan, pour $p = 2, 3$. Il faut bien comprendre qu'un point a ses coordonnées modulo p. Les droites vectorielles peuvent être vues comme des droites passant par l'origine, située en bas à gauche.

Notons que la variété de drapeaux sur les corps finis est en lien direct avec la théorie de Sylow.

1.4. Proposition. *Soit q une puissance d'un nombre premier p et $n \in \mathbb{N}^*$. Les p-Sylow de $\mathrm{GL}_n(\mathbb{F}_q)$ sont isomorphes au groupe TSU_n des matrices triangulaires supérieures avec des 1 sur la diagonale. L'ensemble de ces p-Sylow est en bijection naturelle[2] avec les drapeaux complets de $\mathbb{F}_q^n$.*

Démonstration. Soit $G = \mathrm{GL}_n(\mathbb{F}_q)$. Par la proposition 1.1, l'ordre du groupe G est $q^{n(n-1)/2}(q^n-1)(q^{n-1}-1)\times\cdots\times(q-1)$. On en déduit que les p-Sylow de G sont d'ordre $q^{n(n-1)/2}$.

La partie TSU_n est clairement un sous-groupe de $\mathrm{GL}_n(\mathbb{F}_q)$ d'ordre $q^{n(n-1)/2}$. C'est donc bien un q-Sylow de G.

On sait qu'un groupe fini G agit transitivement par conjugaison sur l'ensemble $\mathscr{S}$ de ses q-Sylow. L'ensemble $\mathscr{S}$ est donc en bijection avec G/N, où N est le normalisateur d'un Sylow.

Appelons donc N le normalisateur de TSU_n et montrons qu'il s'agit du sous-groupe TS_n des matrices triangulaires supérieures de G.

On a clairement $TS_n \subset N$, puisque si $b \in TS_n$, $u \in TSU_n$, alors bub^{-1} est triangulaire supérieure et sa seule valeur propre est 1.

Réciproquement, pour $m \in N$, établissons que $m \in TS_n$. Tout d'abord, soit $(e_i)_{1\leqslant i\leqslant n}$ la base canonique de $\mathbb{F}_q^n$; montrons alors que e_1 est vecteur propre pour m. Pour cela, on remarque que la droite $\mathbb{F}_q e_1$ est l'intersection

[2]La bijection est aussi naturelle que la base canonique...

des espaces propres des éléments de TSU_n. Soit $u \in TSU_n$; alors, comme m normalise TSU_n, il existe v dans TSU_n tel que $u\big(m(e_1)\big) = m\big(v(e_1)\big) = m(e_1)$. Conclusion : $m(e_1)$ appartient à l'intersection des espaces propres des éléments de TSU_n, et donc est de la forme ae_1, $a \neq 0$. En quotientant l'espace $\mathbb{F}_q^n$ par $\mathbb{F}_q e_1$ (on vient de voir que m passe au quotient !), on obtient par récurrence que m est triangulaire supérieure.

Ainsi, $\mathscr{S}$ est en bijection avec G/TS_n, qui est lui-même en bijection avec la variété de drapeaux. □

1.5. Remarque. On vérifie que le nombre de q-Sylow est bien congru à 1 modulo q.

2. Décomposition en cellules et dénombrement

On notera $\mathbb{A}^n$ (ou $\mathbb{A}^n(\mathbb{K})$ s'il y a ambiguïté [3]) l'espace affine $\mathbb{K}^n$.

Nous profitons du calcul du cardinal de l'espace projectif pour la relier à la décomposition en cellules du paragraphe IV-3. Ce qui vient peut toutefois se lire de façon indépendante. Il suffit de se rappeler qu'une décomposition en cellules d'un espace topologique est une partition dont les parts, appelées cellules, sont homéomorphes à des espaces affines ; on impose de plus que l'adhérence d'une cellule soit une réunion de cellules.

Cette notion est moins contraignante que ce qui est communément appelé « décomposition cellulaire », notion définie dans le livre [16] de Claude Godbillon.

La formule $(q^n - 1)/(q - 1) = q^{n-1} + \cdots + 1$ suggère que l'espace projectif $\mathbb{P}^{n-1}$ se décompose en une réunion disjointe d'espaces vectoriels. De façon plus naturelle, nous allons voir qu'il s'agit en fait d'espaces affines appelées cellules. Pour donner un exemple simple, il arrive fréquemment que l'on scinde en deux l'ensemble des droites vectorielles du plan : celles qui sont de la forme $y = ax$ et la droite « verticale » $x = 0$. C'est une illustration de la décomposition en cellules $\mathbb{P}^1(\mathbb{F}_q) = \mathbb{A}^1 \sqcup \mathbb{A}^0$, qui incarne la formule $(q^2 - 1)/(q - 1) = q^1 + 1$.

Ces décompositions possèdent un double intérêt. D'une part, elles vont permettre de voir que la géométrie projective « prolonge » (il vaut mieux dire « complète ») la géométrie affine. D'autre part, sur $\mathbb{R}$ et $\mathbb{C}$, les implications topologiques seront importantes, puisque l'on sera capable de décomposer l'espace topologique complexe (au sens de compliqué) $\mathbb{P}^n$ en union d'espaces topologiques simplissimes (un espace affine est simplement connexe).

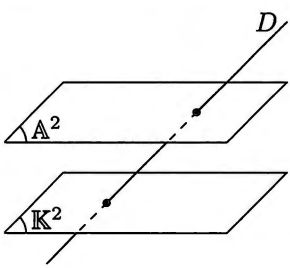
[3] On préférera plutôt, dans ce cas, l'abus de notation $\mathbb{K}^n$.

2.1. Proposition. *Soit $\mathbb{K}$ un corps et n un entier naturel. On a une décomposition en cellules :*

$$\mathbb{P}^n(\mathbb{K}) \simeq \mathbb{A}^n \sqcup \mathbb{P}^{n-1}(\mathbb{K}) \simeq \mathbb{A}^n \sqcup \mathbb{A}^{n-1} \sqcup \cdots \sqcup \mathbb{A}^1 \sqcup \mathbb{A}^0.$$

Nous laissons le soin au lecteur de compléter la preuve pour tout n, une preuve plus générale de décomposition de la grassmannienne (n'oublions pas que l'espace projectif n'est qu'un cas particulier de grassmannienne $\mathrm{Gr}_{m,n}$ pour $m = 1$) étant donnée en proposition IV-3.2.4. Voici tout de même une illustration de ce résultat pour les cas $n = 1$, $n = 2$.

2.2. Exemple. Dans l'exemple qui suit, on réalise $\mathbb{A}^2$ dans $\mathbb{K}^3$ comme étant l'hyperplan d'équation $z_3 = 1$, où z_3 désigne la dernière coordonnée de $\mathbb{K}^3$. L'ensemble $\mathbb{P}^2$ se divise alors en deux parties disjointes, celle des droites coupant $\mathbb{A}^2$, et celle des droites ne le coupant pas, assimilée à l'ensemble des droites de $\mathbb{K}^3$.



$$\begin{aligned}\mathbb{P}^2(\mathbb{K}) &= \{\text{droites de } \mathbb{K}^3\} = \mathbb{A}^2 \sqcup \mathbb{P}^1(\mathbb{K}) \\ &= \{\text{droites coupant } \mathbb{A}^2\} \sqcup \{\text{droites de } \mathbb{K}^2\}\end{aligned}$$

Figure 2.1. *Plan projectif* $\mathbb{P}^2(\mathbb{K})$

En conclusion, l'espace projectif est la réunion disjointe de l'espace affine correspondant et d'un espace projectif de taille inférieure, appelé *hyperplan à l'infini*. On pourra donc assimiler les droites d'un espace vectoriel à un point d'un espace affine. Nous verrons dans le chapitre suivant une décomposition en cellules de la grassmannienne.

3. Applications aux isomorphismes exceptionnels de groupes finis

On sait que les groupes finis simples (qui n'ont aucun sous-groupe distingué) ont été classés à isomorphisme près par à un tour de force acharné de mathématiciens du vingtième siècle. On rappelle que parmi les groupes

finis simples non abéliens, on peut compter deux séries : les groupes alternés $\mathfrak{A}_n$, $n \geqslant 5$, et les groupes $\mathrm{PSL}_n(\mathbb{F}_q)$, $(n,q) \neq (2,2), (2,3)$, voir [29], théorèmes I 8.1 et IV 4.1.

Il est naturel de se demander si les groupes de ces deux séries peuvent accidentellement être isomorphes (ou même s'il existe des isomorphismes à l'intérieur d'une telle série). La réponse est : « en général non, mais exceptionnellement, oui ». Cela amène à la notion d'isomorphismes exceptionnels. Cette notion possède aussi une importance pratique pour le mathématicien de terrain : si l'on sait par exemple que $\mathrm{PSL}_2(\mathbb{F}_4)$ est isomorphe à $\mathfrak{A}_5$ lui-même isomorphe à $\mathrm{PSL}_2(\mathbb{F}_5)$, on a alors trois façons distinctes d'approcher ce groupe, et on le connaîtra d'autant trois fois mieux. Ces isomorphismes se démontrent, non pas à la main, par construction par générateurs et relations, mais via une action de nature géométrique. C'est ce qui les rend si passionnants : ils sont à chaque fois la manifestation abstraite d'un phénomène géométrique subtil. Ils sont d'ailleurs si passionnants que l'on ne finit par s'intéresser qu'à eux, ce qui en fait oublier leur caractère exceptionnel...

On rappelle le cadre général. Soit G un groupe agissant sur un ensemble X de cardinal n. Alors, on a un morphisme :

$$\begin{aligned} \varphi : G &\longrightarrow \mathfrak{S}_X \simeq \mathfrak{S}_n \\ g &\longmapsto (\varphi_g : X \to X), \end{aligned}$$

d'où un isomorphisme du quotient $G/\operatorname{Ker}(\varphi)$ avec le sous-groupe $\operatorname{Im}\varphi$ de $\mathfrak{S}_n$.

On rappelle aussi que l'action naturelle du groupe linéaire $\mathrm{GL}_n(\mathbb{K})$ sur l'ensemble $\mathbb{P}^{n-1}(\mathbb{K})$ des droites vectorielles de $\mathbb{K}^n$ admet pour noyau le sous-groupe $\operatorname{Ker}\varphi = \mathbb{K}^*$ des homothéties (un endomorphisme qui stabilise toutes les droites est une homothétie), d'où l'on déduit une action fidèle du groupe $\mathrm{PGL}_n(\mathbb{K}) := \mathrm{GL}_n(\mathbb{K})/\mathbb{K}^*$, qui apparaît ainsi comme sous-groupe de bijections de $\mathbb{P}^{n-1}(\mathbb{K})$ dans lui-même.

Certaines permutations de droites peuvent ainsi être réalisées par une transformation projective et d'autres non. C'est ce que les isomorphismes exceptionnels du chapitre essaient de déterminer.

Voici une liste d'isomorphismes exceptionnels classiques de groupes finis.

3.1. Proposition.

(i) $\mathrm{GL}_2(\mathbb{F}_2) \simeq \mathrm{SL}_2(\mathbb{F}_2) \simeq \mathrm{PSL}_2(\mathbb{F}_2) \simeq \mathrm{PGL}_2(\mathbb{F}_2) \simeq \mathfrak{S}_3$;

(ii) $\mathrm{PSL}_2(\mathbb{F}_3) \simeq \mathfrak{A}_4$, $\mathrm{PGL}_2(\mathbb{F}_3) \simeq \mathfrak{S}_4$;

(iii) $\mathrm{PGL}_2(\mathbb{F}_4) = \mathrm{PSL}_2(\mathbb{F}_4) \simeq \mathfrak{A}_5$;

(iv) $\mathrm{PSL}_2(\mathbb{F}_5) \simeq \mathfrak{A}_5$, $\mathrm{PGL}_2(\mathbb{F}_5) \simeq \mathfrak{S}_5$.

LU PRODUIT DE L'ANNÉE

Démonstration. (i) Comme $\mathbb{F}_2^* = \{1\}$, on a d'une part : $\mathrm{PGL}_2(\mathbb{F}_2) = \mathrm{GL}_2(\mathbb{F}_2)/\mathbb{F}_2^* = \mathrm{GL}_2(\mathbb{F}_2)$ et, de même : $\mathrm{PSL}_2(\mathbb{F}_2) = \mathrm{SL}_2(\mathbb{F}_2)$. D'autre part, le déterminant est le morphisme trivial, d'où : $\mathrm{GL}_2(\mathbb{F}_2) = \mathrm{SL}_2(\mathbb{F}_2)$.

Par ailleurs, le groupe $\mathrm{GL}_2(\mathbb{F}_2)$ agit sur $\mathbb{P}^1(\mathbb{F}_2)$, qui a trois éléments, d'où une injection : $\mathrm{PGL}_2(\mathbb{F}_2) \hookrightarrow \mathfrak{S}_{\mathbb{P}^1(\mathbb{F}_2)} = \mathfrak{S}_3$. Comme on a :

$$|\mathrm{PGL}_2(\mathbb{F}_2)| = \frac{(2^2-1)(2^2-2)}{2-1} = 6 = |\mathfrak{S}_3|,$$

on en conclut que $\mathrm{PGL}_2(\mathbb{F}_2) \simeq \mathfrak{S}_3$.

(ii) Le groupe $\mathrm{GL}_2(\mathbb{F}_3)$ agit sur $\mathbb{P}^1(\mathbb{F}_3)$, qui a trois éléments, d'où une injection : $\mathrm{PGL}_2(\mathbb{F}_3) \hookrightarrow \mathfrak{S}_{\mathbb{P}^1(\mathbb{F}_3)} = \mathfrak{S}_4$. On calcule :

$$|\mathrm{PGL}_2(\mathbb{F}_3)| = \frac{(3^2-1)(3^2-3)}{3-1} = 24 = |\mathfrak{S}_4|,$$

et l'on en conclut que $\mathrm{PGL}_2(\mathbb{F}_3) \simeq \mathfrak{S}_4$.

Enfin, on a : $\mathrm{PSL}_2(\mathbb{F}_3) \simeq \mathfrak{A}_4$, car $\mathfrak{A}_4$ est le seul sous groupe d'indice $2 = |\mathbb{F}_3^*|$ dans $\mathfrak{S}_4$.

(iii) Mêmes arguments. Laissé en exercice.

(iv) Le groupe $\mathrm{GL}_2(\mathbb{F}_5)$ agit sur $\mathbb{P}^1(\mathbb{F}_5)$, qui a six éléments, ce qui fournit une injection $\mathrm{PGL}_2(\mathbb{F}_5) \hookrightarrow \mathfrak{S}_{\mathbb{P}^1(\mathbb{F}_5)} = \mathfrak{S}_6$. On calcule alors : $|\mathrm{PGL}_2(\mathbb{F}_5)| = (5^2-1)(5^2-5)/(5-1) = 120$.

Ainsi : $[\mathfrak{S}_6 : \mathrm{PGL}_2(\mathbb{F}_5)] = 6$, on en conclut que $\mathrm{PGL}_2(\mathbb{F}_5) \simeq \mathfrak{S}_5$ car, quel que soit l'entier n, tout sous-groupe d'indice n dans $\mathfrak{S}_n$ est isomorphe à $\mathfrak{S}_{n-1}$, voir [29], corollaire I- 8.6. □

3.2. Remarque. On a donc une action de $\mathrm{PGL}_2(\mathbb{F}_5) \subset \mathfrak{S}_5$ sur les six droites de $\mathbb{P}^1(\mathbb{F}_5)$, donc un plongement de $\mathfrak{S}_5$ dans $\mathfrak{S}_6$, qui ne stabilise aucun élément. Via la proposition VIII-1.4 et le constat qu'une droite dans un plan « est » un drapeau complet, on identifie cette action à l'action sur les 5-Sylow de $\mathfrak{S}_5$, qui sont au nombre de 6.

On peut déduire de cette action l'existence d'un automorphisme extérieur dans $\mathrm{Aut}(\mathfrak{S}_6)$, voir [29] ou l'exercice VIII-A.8.

3.3. Remarque. L'action du groupe linéaire sur l'espace projectif est valable pour tout corps $\mathbb{K}$: le groupe $\mathrm{GL}_{n+1}(\mathbb{K})$ agit sur $\mathbb{P}^n(\mathbb{K})$ et c'est même une action transitive (exercice !) de stabilisateur $P_{1,n}(\mathbb{K}) = \begin{pmatrix} \mathbb{K}^* & \mathbb{K} \\ 0 & \mathrm{GL}_n(\mathbb{K}) \end{pmatrix}$, un sous-groupe parabolique.

Quand on est sur $\mathbb{R}$ ou $\mathbb{C}$, le théorème d'homéomorphisme donne, une fois ses hypothèses bien vérifiées, $\mathbb{P}^n(\mathbb{K}) \cong \mathrm{GL}_{n+1}(\mathbb{K})/P_{1,n}(\mathbb{K})$.

3.4. Remarque. D'autres isomorphismes exceptionnels de groupes finis simples sont dispersés au second tome, chapitres V et XII. Nous introduirons à ce moment-là des méthodes plus générales (et peut-être plus canoniques) pour les construire.

Dans le même ordre d'idée, voici un isomorphisme qui en a étonné plus d'un. Pourtant, cette proposition ne fait que remarquer que le « cercle est cyclique » (le scoop !). La seconde partie de la preuve est instructive, car on y utilise un dessin sur le plan réel pour raisonner sur de la géométrie de corps finis.

3.5. Proposition. *Soit q une puissance d'un nombre premier impair. On a un isomorphisme :*

$$\mathrm{SO}_2(\mathbb{F}_q) \simeq \begin{cases} \mathbb{Z}/(q-1)\mathbb{Z} & \text{si } -1 \text{ est un carré de } \mathbb{F}_q^*, \\ \mathbb{Z}/(q+1)\mathbb{Z} & \text{si } -1 \text{ n'est pas un carré de } \mathbb{F}_q^*. \end{cases}$$

Démonstration. Le groupe se décrit par :

$$\begin{aligned} \mathrm{SO}_2(\mathbb{F}_q) &= \left\{ A \in \mathrm{GL}_2(\mathbb{F}_q) : \det A = 1,\ {}^tAA = \mathrm{I}_2 \right\} \\ &= \left\{ \begin{pmatrix} a & c \\ b & d \end{pmatrix} : ad - bc = 1,\ a^2 + b^2 = 1,\ c^2 + d^2 = 1,\ ac + bd = 0 \right\}. \end{aligned}$$

Fixons (a, b) tel que $a^2 + b^2 = 1$. Les équations

$$\begin{cases} ac + bd = 0 \\ -bc + ad = 1 \end{cases}$$

forment un système de déterminant : $a^2 + b^2 = 1$. La solution évidente $(c, d) = (-b, a)$ est donc la seule, et l'on a bien : $c^2 + d^2 = 1$.

En notant $\mathbf{S}^1(\mathbb{F}_q)$ le « cercle unité », c'est-à-dire l'ensemble des (a, b) tels que $a^2 + b^2 = 1$, on a donc une élégante bijection :

$$\mathbf{S}^1(\mathbb{F}_q) \longrightarrow \mathrm{SO}_2(\mathbb{F}_q), \quad (a, b) \longmapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Il s'agit donc de compter les points du cercle.

Premier cas : $-1 \in \mathbb{F}_q^{*2}$. Soit $\omega \in \mathbb{F}_q^*$ tel que $-1 = \omega^2$.

Pour résoudre $a^2 + b^2 = 1$, on peut factoriser : $a^2 + b^2 = (a - \omega b)(a + \omega b)$ et faire un changement de variable, justifié car q est impair ($2 \neq 0$) :

$$\begin{cases} x = a + \omega b \\ y = a - \omega b \end{cases} \iff \begin{cases} a = \dfrac{x + y}{2} \\ b = \dfrac{x - y}{2\omega}. \end{cases}$$

Et, par ce changement de variable, on obtient :

$$|\mathrm{SO}_2(\mathbb{F}_q)| = |\mathbf{S}^1(\mathbb{F}_q)| = |\{(x,y) \in \mathbb{F}_q^2,\ xy = 1\}|.$$

Le choix de $x \in \mathbb{F}_q^*$ détermine un et un seul couple, d'où : $|\mathrm{SO}_2(\mathbb{F}_q)| = q-1$.

De plus, on vérifie que

$$\psi : \mathrm{SO}_2(\mathbb{F}_q) \longrightarrow \mathbb{F}_q^*, \quad \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \longmapsto a + \omega b$$

est un morphisme (vérification directe, que l'on peut aussi comparer avec son analogue dans le corps des complexes). Il est injectif : en effet, si $\psi(A) = 1$ alors $x = a + \omega b = 1$, puis $y = a - \omega b = (a^2 + b^2)/(a + \omega b) = 1$, enfin, $a = 1$ et $b = 0$.

Par égalité des cardinaux, ψ est un isomorphisme de $\mathrm{SO}_2(\mathbb{F}_q)$ sur le groupe cyclique $\mathbb{F}_q^*$. Bilan : $\mathrm{SO}_2(\mathbb{F}_q) \simeq \mathbb{Z}/(q-1)\mathbb{Z}$.

Deuxième cas : $-1 \notin \mathbb{F}_q^{*2}$.

Raisonnons par analogie avec les réels (quoi qu'il en soit, on raisonne toujours sur des dessins faux !). Voir la figure 3.1.

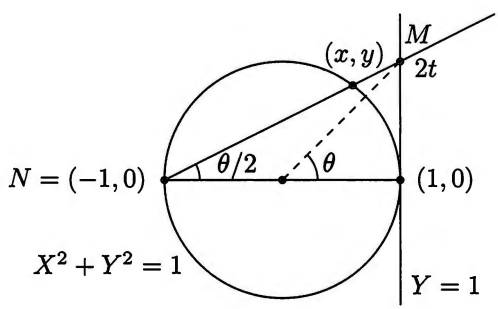


Figure 3.1. *Paramétrage du cercle unité privé d'un point par une droite*

Soit $N = (-1, 0)$ dans le plan $\mathbb{F}_q^2$, dont on note (X, Y) les coordonnées. Soit $t \in \mathbb{F}_q$, on note $M = (1, 2t)$. La droite (NM) coupe le cercle $\mathbf{S}^1(\mathbb{F}_q)$ en N et en un deuxième point $p(t)$. (Sur les réels, cela revient à paramétrer le point (x, y) par t, c'est-à-dire, avec le théorème de l'angle inscrit : $t = \tan(\theta/2)$.)

En effet, même si c'est une propriété générale des coniques, on le vérifie à la main : (NM) admet pour équation $y = t(x+1)$ (on utilise $2 \neq 0$), d'où l'équation aux abscisses : $(1+t^2)x^2 + 2t^2x + t^2 - 1 = 0$. C'est une équation de degré 2, car $1 + t^2 \neq 0$ (-1 n'est pas un carré). La solution évidente -1

étant supprimée, l'autre s'obtient sans calcul par le produit des racines :

$$x = \frac{1-t^2}{1+t^2}, \quad y = t(x+1) = \frac{2t}{1+t^2}.$$

(Sur les réels, on exprime ainsi $\cos\theta$ et $\sin\theta$ en fonction de $\tan(\theta/2)$.)

Inversement, pour tout point $M' = (x, y)$ de $\mathbf{S}^1(\mathbb{F}_q)$ autre que N, on a : $x \neq 1$, donc la droite (NM') coupe la droite $Y = 1$ en un point. Ainsi, p établit une bijection de $\mathbb{F}_q$ sur $\mathbf{S}^1(\mathbb{F}_q) \setminus \{N\}$. On en déduit :

$$|\,\mathrm{SO}_2(\mathbb{F}_q)| = |\mathbf{S}^1(\mathbb{F}_q)| = q+1.$$

Pour montrer que $\mathrm{SO}_2(\mathbb{F}_q)$ est cyclique dans ce cas, la preuve est analogue au cas précédent, sauf que l'on injecte $\mathrm{SO}_2(\mathbb{F}_q)$ dans $\mathbb{F}_{q^2}^*$, extension dans laquelle -1 possède une racine carrée ω. On utilise alors le fait que le groupe multiplicatif d'un corps fini quelconque est cyclique et que tout sous-groupe d'un groupe cyclique est lui-même cyclique. □

3.6. Remarque. On peut reformuler la fin de la preuve de façon un peu plus conceptuelle. Introduisons les algèbres sur $\mathbb{F}_q$ suivantes[4] :

$$E = \mathbb{F}_q[X]/(X^2+1) \quad \text{et} \quad C = \left\{ A_{a,b} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix},\ a, b \in \mathbb{F}_q \right\}.$$

Notons $\mathbf{i} \in E$ la classe de X modulo X^2+1. L'application $E \to C$, $a+b\mathbf{i} \mapsto A_{a,b}$ est un isomorphisme : en effet, $A_{a,b}$ est la matrice de la multiplication par $a+b\mathbf{i}$ dans le $\mathbb{F}_q$-espace de dimension 2 qu'est E. (Remplaçant $\mathbb{F}_q$ par $\mathbb{R}$, on retrouve la réalisation matricielle des complexes du § VII-1.2.)

Revenons à $\mathrm{SO}_2(\mathbb{F}_q)$: c'est un sous-groupe du groupe $C^\times$ des inversibles de C, donc un sous-groupe des inversibles $E^\times$ de E. La multiplicativité du déterminant exprime $\mathrm{SO}_2(\mathbb{F}_q)$ comme le noyau de det : $C^\times \to \mathbb{F}_q^*$.

Premier cas : -1 est un carré, disons $-1 = \omega^2$. Alors, par le lemme chinois, on a un isomorphisme : $E \simeq \mathbb{F}_q \times \mathbb{F}_q$, $a+b\mathbf{i} \mapsto (x, y)$, où $x = a+b\omega$ et $y = a-b\omega$ comme avant. Alors, $\mathrm{SO}_2(\mathbb{F}_q) \simeq \{(x, x^{-1}),\ x \in \mathbb{F}_q^*\} \simeq \mathbb{F}_q^*$.

Deuxième cas : -1 n'est pas un carré. Alors, $E = \mathbb{F}_{q^2}$. Le déterminant de $A_{a,b}$ est la norme a^2+b^2 de $a+\mathbf{i}b$, où la norme N est une application multiplicative et surjective[5] $\mathbb{F}_{q^2} \to \mathbb{F}_q$. Par suite, $\mathrm{SO}_2(\mathbb{F}_q)$, qui est le noyau de $N : \mathbb{F}_{q^2}^* \twoheadrightarrow \mathbb{F}_q^*$, est un groupe cyclique de cardinal : $(q^2-1)/(q-1) = q+1$.

[4] La lettre E est censée évoquer une extension, C rappeler les nombres complexes.

[5] Tout élément de $\mathbb{F}_q^*$ admet une racine dans $\mathbb{F}_{q^2}$. Dans notre cas, si z n'est pas un carré de $\mathbb{F}_q$, alors $-z$ admet en effet une racine a (multiplicativité du symbole de Legendre) donc z en a une, $a\mathbf{i}$, dans $\mathbb{F}_{q^2}$.

4. Développement. Un non-isomorphisme exceptionnel

All work and no play makes Jack a dull boy.

The Shining, Stanley Kubrick, 1980.

Peut-il arriver que deux groupes simples non isomorphes aient le même ordre ? C'est rare lorsque les cardinaux sont « petits », mais un exercice du chapitre IV de [29] en donne un premier exemple à l'ordre 20 160. Nous allons le montrer.

4.1. Proposition. *Les groupes* $\mathrm{PSL}(4,\mathbb{F}_2)$ *et* $\mathrm{PSL}(3,\mathbb{F}_4)$ *sont deux groupes simples non isomorphes de même ordre* 20 160.

Démonstration. Nous avons déjà admis que ces deux groupes sont simples, voir [29].

Remarquons tout d'abord que si une homothétie de rapport λ est dans SL_2, alors $\lambda^2 = 1$, et donc en caractéristique 2, il s'agit d'une homothétie de rapport 1 : c'est l'identité. Posons donc $G := \mathrm{PSL}(4,\mathbb{F}_2) = \mathrm{SL}(4,\mathbb{F}_2)$ et $H = \mathrm{PSL}(3,\mathbb{F}_4) = \mathrm{SL}(3,\mathbb{F}_4)$. Comme 1 est le seul élément non nul de $\mathbb{F}_2$, on a aussi : $G = \mathrm{GL}(4,\mathbb{F}_2)$.

L'ordre du groupe linéaire G est donné par la proposition 1.1 et vaut ainsi $(2^4-1)(2^4-2)(2^4-2^2)(2^4-2^3)$ et celui de H $(4^3-1)(4^3-4)(4^3-4^2)/(4-1)$, et l'on vérifie ainsi la coïncidence numérique.

L'idée de la preuve est d'utiliser la réduction des endomorphismes pour montrer que G et H ne sont pas isomorphes, en calculant le nombre respectif de classes de conjugaison de leurs éléments d'ordre 2. En effet, l'ensemble des éléments d'ordre 2 d'un groupe est stable par conjugaison et si deux groupes finis sont isomorphes alors un isomorphisme enverra l'ensemble des éléments d'ordre 2 de l'un vers l'ensemble des éléments d'ordre 2 de l'autre et une classe de conjugaison de l'un vers une classe de conjugaison de l'autre. Par conséquent, si l'on montre que le nombre de classes de conjugaison d'éléments d'ordre 2 de G est différent du nombre de classes de conjugaison de H, on aura prouvé que G et H ne sont pas isomorphes.

Un élément d'ordre 2 de GL_n a pour polynôme minimal $X^2-1 = (X-1)^2$ en caractéristique 2. Sa décomposition de Dunford est donc $\mathrm{I}_n + N$, où N est nilpotente d'indice 2. Par le théorème III-2.5.2, le nombre de classes de conjugaison est donc donné par le nombre de partitions de n commençant par 2. Pour $n = 4$, les partitions sont $2 \geqslant 2$ et $2 \geqslant 1 \geqslant 1$. Donc, G possède deux telles classes de conjugaison.

De même, $\mathrm{GL}(3,\mathbb{F}_4)$ possède une seule classe de conjugaison d'éléments d'ordre 2 correspondant à la partition $2 \geqslant 1$. De plus, cette classe de conjugaison est incluse dans $H = \mathrm{SL}(3,\mathbb{F}_4)$ (1 est seule valeur propre).

Mais attention ! Les classes de conjugaison pour un groupe G incluses dans un sous-groupe H peuvent se scinder en plusieurs classes de conjugaison pour H.

Montrons que H possède une seule classe de conjugaison d'ordre 2, ce qui achèvera la preuve. Pour cela, il suffit de montrer que deux éléments d'ordre 2 conjugués pour $\mathrm{GL}(3,\mathbb{F}_4)$ sont aussi conjugués pour $\mathrm{SL}(3,\mathbb{F}_4)$. Soit

$$J := \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Par transitivité, il suffit de montrer que si un élément est $\mathrm{GL}(3,\mathbb{F}_4)$-conjugué à J, alors il est $\mathrm{SL}(3,\mathbb{F}_4)$-conjugué à J.

Soit donc $A \in H$ et $P \in \mathrm{GL}(3,\mathbb{F}_4)$ tels que $A = PJP^{-1}$. Posons alors

$$Q = P \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \det P^{-1} \end{pmatrix}.$$

On a bien : $Q \in H$ et $A = QJQ^{-1}$. □

4.2. Remarque. Le groupe $\mathrm{PSL}(4,\mathbb{F}_2)$ est également isomorphe au groupe alterné $\mathfrak{A}_8$. Nous donnerons une idée de la preuve au chapitre XII du second tome.

4.3. Remarque. Pour trouver un autre non-isomorphisme de groupes finis simples de même cardinal, il faut aller le chercher à l'ordre 4 585 351 680.

A. Exercices

Dans les exercices qui suivent, on fixe un entier naturel n et un corps fini $\mathbb{F}_q$ de cardinal q.

A.1. Exercice

1. *Combien y-a-t-il de systèmes de n vecteurs indépendants dans F_q^m ?*
   Faire agir le groupe linéaire sur l'ensemble de ces systèmes. On obtient dans un premier temps $|\mathrm{GL}_m(\mathbb{F}_q)|/q^{n(m-n)}|\,\mathrm{GL}_{m-n}(\mathbb{F}_q)|$.

2. *En déduire le nombre d'applications linéaires surjectives de $\mathbb{F}_q^m$ dans $\mathbb{F}_q^n$.*
   Utiliser la question précédente pour trouver le nombre d'applications injectives de $\mathbb{F}_q^n$ dans $\mathbb{F}_q^m$. Puis, utiliser (astuce !) la transposition.

A.2. Exercice. *Calculer le cardinal de l'orbite d'une matrice de Jordan indécomposable sur un corps fini pour la conjugaison.*

Le stabilisateur d'une matrice pour l'action de conjugaison est égal à son commutant. Le commutant d'une matrice de Jordan indécomposable est égal au sous-espace des polynômes en cette matrice, voir exercice III-D.15.

A.3. Exercice. *Calculer le cardinal de l'orbite d'une matrice diagonalisable sur $\mathbb{F}_q$ pour la conjugaison (tenir compte de la multiplicité des valeurs propres).*

Montrer tout d'abord que le commutant d'une matrice diagonale est isomorphe à $\prod_i \mathrm{GL}_{n_i}(\mathbb{F}_q)$, où les n_i sont les multiplicités des valeurs propres.

A.4. Exercice

1. *Combien y a-t-il de matrices carrées $n\times n$ de rang r sur $\mathbb{F}_q$?*
   Déjà vu dans l'exercice I-3.6.
2. *Combien y-a-t-il de matrices carrées $n\times n$ diagonalisables sur $\mathbb{F}_q$ à valeurs propres distinctes ?*

A.5. Exercice (Nombre de matrices diagonalisables de $\mathscr{M}_n(\mathbb{F}_q)$) *Soit n et k deux entiers. On fixe k entiers $n_1, n_2, \ldots, n_k$ de n tels que $\sum_{i=1}^k n_i = n$.*

1. *Montrer que le nombre de familles $(E_1, \cdots, E_k)$ de sous-espaces tels que*
$$E_1 \oplus \cdots \oplus E_k = \mathbb{F}_q^n, \quad \dim E_i = n_i \quad (1 \leqslant i \leqslant k),$$
   est égal à
$$|\mathrm{GL}_n(\mathbb{F}_q)| \Big/ \prod_{i=1}^k |\mathrm{GL}_{n_i}(\mathbb{F}_q)|.$$
   On fait agir de façon naturelle $\mathrm{GL}_n(\mathbb{F}_q)$ sur l'ensemble de ses k-uplets, on montre que l'action est transitive et on regarde le stabilisateur.
   On rappelle qu'une matrice A de $\mathscr{M}_n(\mathbb{F}_q)$ est diagonalisable si et seulement si $A^q = A$ (exercice III-D.5*).*

2. En déduire que le nombre de matrices diagonalisables de $\mathcal{M}_n(\mathbb{F}_q)$ est :

$$\sum_{n_1+\cdots+n_q=n, n_i\geqslant 0} \frac{|\operatorname{GL}_n(\mathbb{F}_q)|}{\prod_{i=1}^{q} |\operatorname{GL}_{n_i}(\mathbb{F}_q)|}.$$

Penser au lemme des noyaux sur la matrice A, annulée par $X^q - X$. On construira une bijection en assignant à la matrice diagonalisable A le q-uplet de sous-espaces $(E_\zeta,\ \zeta \in \mathbb{F}_q)$, où E_ζ est le sous-espace propre de A pour la valeur propre ζ.

Cet exercice est inspiré d'un texte écrit par Mathieu Romagny pour les agrégatifs sur sa page personnelle.

A.6. Exercice. *On fixe un sous-espace vectoriel F de dimension m dans un espace E de dimension n sur $\mathbb{F}_q$. En faisant agir le stabilisateur de F dans $\operatorname{GL}(E)$, montrer les formules suivantes :*

$$\#\{G,\ F\oplus G = E\} = q^{m(n-m)},$$

$$\#\{G,\ F+G=E, \dim F\cap G = k\} = \begin{bmatrix} m \\ k \end{bmatrix}_q q^{(m-k)(n-m)},$$

$$\#\{G,\ \dim G = h, \dim F\cap G = k\} = \begin{bmatrix} m \\ k \end{bmatrix}_q \begin{bmatrix} n-m \\ h-k \end{bmatrix}_q q^{m(n-m)-k(h-k)},$$

où le symbole $\begin{bmatrix} n \\ m \end{bmatrix}_q$ désigne le cardinal de $|\operatorname{Gr}_{m,n}(\mathbb{F}_q)|$, soit

$$\begin{bmatrix} n \\ m \end{bmatrix}_q = \frac{(q^n-1)(q^{n-1}-1)\times\cdots\times(q^{n-m+1}-1)}{(q^m-1)(q^{m-1}-1)\times\cdots\times(q-1)}.$$

A.7. Exercice. *Combien y a -t-il de formes quadratiques de rang fixé sur l'espace $(\mathbb{F}_q)^n$?*

On calculera le nombre de matrices symétriques de rang r, dont on sait qu'elles forment deux orbites pour $r \neq 0$ pour l'action de congruence de $\operatorname{GL}_n(\mathbb{F}_q)$, voir l'exercice V-1.4.

A.8. Exercice (Quelques sous-groupes du groupe symétrique)

1. Montrer que pour $n \geqslant 5$, les seuls sous-groupes distingués de $\mathfrak{S}_n$ sont $\{1\}, \mathfrak{A}_n$ et $\mathfrak{S}_n$.

On sait que $\mathfrak{A}_n$ est simple dans ce cas.

2. *On va montrer la propriété suivante.*

   Soit H un sous-groupe d'indice n de $G := \mathfrak{S}_n$. Alors, H est isomorphe à $\mathfrak{S}_{n-1}$.

   (a) Montrer l'énoncé pour $n \leqslant 4$.

   C'est du cas par cas !

   (b) Supposons que $n \geqslant 5$. Considérons l'opération du groupe G par translation à gauche sur le quotient G/H. On a donc un homomorphisme $\varphi : \mathfrak{S}_n \longrightarrow \mathfrak{S}(G/H) \cong \mathfrak{S}_n$. Montrer que $\operatorname{Ker}\varphi = \bigcap_{g\in G} gHg^{-1}$. En déduire que φ est injective.

   (c) Montrer que le stabilisateur de $\overline{1} = H$ est $\varphi(H)$. Conclure.

A.9. Exercice (Groupe orthogonal $\mathrm{O}(2, \mathbb{F}_q)$ pour toute caractéristique !)

Soit q une puissance d'un nombre premier.

1. *On suppose q impair. Montrer que $\mathrm{O}_2(\mathbb{F}_q)$ est le produit semi-direct de $\mathbb{Z}/2\mathbb{Z}$ par le groupe cyclique $\mathrm{SO}_2(\mathbb{F}_q)$.*
   Le plus simple est de trouver une section au morphisme déterminant $\mathrm{O}_2(\mathbb{F}_q) \to \{\pm 1\}$ et d'appliquer le corollaire II-5.3.4.
2. *En déduire que $\mathrm{O}(2, \mathbb{F}_q)$ est isomorphe au groupe diédral.*
   Utiliser la proposition 3.5 et faire agir une symétrie orthogonale particulière par conjugaison sur les rotations.
3. *Soit $\mathbb{K}$ un corps de caractéristique 2 et M une matrice de $\mathscr{M}_2(\mathbb{K})$:*

$$M = \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

   (a) Montrer l'équivalence :

$$M \in \mathrm{O}_2(\mathbb{K}) \iff b = c = a + 1 = d + 1.$$

   On traduira l'égalité ${}^tMM = \mathrm{I}_2$ par un système et l'on utilisera l'égalité $a^2 + b^2 - 1 = (a + b - 1)^2$ (le signe dans « -1 » est ici pour la forme : on rappelle que $-1 = 1$).

   (b) En déduire que toute matrice M de $\mathrm{O}(2, \mathbb{K})$ satisfait à : $M^2 = \mathrm{I}_2$.
   Sans calcul, s'il vous plaît ! La matrice est symétrique...

   (c) En déduire que $O_2(\mathbb{K})$ est abélien.

   (d) On suppose $\mathbb{K}$ de cardinal fini q, et donc par hypothèse, $q = 2^n$ et $\mathbb{K} = \mathbb{F}_q$. Montrer l'isomorphisme :

$$\mathrm{O}(2, \mathbb{F}_q) = \mathrm{SO}(2, \mathbb{F}_q) \simeq (\mathbb{Z}/2\mathbb{Z})^n.$$

   On peut invoquer le théorème de structure des groupes abéliens finis ou remarquer qu'un groupe où tout élément est d'ordre 2 est (abélien et est) un espace vectoriel sur $\mathbb{F}_2$ et invoquer, *once again*, le théorème de la base incomplète.

Chapitre IX

Groupes de Lie classiques et isomorphismes exceptionnels

"I didn't say it would be easy, Neo. I just said it would be the truth."
Matrix, Andy & Larry Wachowski, 1999.

À ce stade, l'utilité d'une structure topologique sur un groupe devrait être bien comprise. Il est temps d'ajouter une structure de variété différentielle.

Pour illustrer l'apport de cette structure, revenons sur l'isomorphisme étudié dans le chapitre sur les quaternions, $\mathrm{SU}(2)/\{1,-1\} \simeq \mathrm{SO}(3)$. On l'a construit à partir du morphisme $\mathrm{SU}(2) \to \mathrm{O}(3)$ fourni par l'action par conjugaison dans les quaternions. Puis, des considérations topologiques ont permis de montrer que l'image était contenue dans la composante neutre, d'où un morphisme $\varphi : \mathrm{SU}(2) \to \mathrm{SO}(3)$. Mais sa surjectivité a été montrée par une méthode *ad hoc*. Une méthode plus simple et plus générale consisterait à montrer que $\mathrm{Im}\,\varphi$ contient un voisinage ouvert de l'identité, ce qui assure bien la surjectivité par connexité de $\mathrm{SO}(3)$. Or, grâce à une structure de variété différentielle sur les groupes $\mathrm{SU}(2)$ et $\mathrm{SO}(3)$, on peut justement vérifier l'existence d'un tel voisinage, via le théorème d'inversion locale.

On trouvera en appendice quelques définitions et résultats utiles de calcul différentiel (sous-variétés, espaces tangents, théorèmes de submersion) qui seront utilisés dans cette leçon.

1. Groupes de Lie classiques

1.1. Définition. On appelle[1] *groupe de Lie* un groupe topologique muni d'une structure de sous-variété de $\mathbb{R}^N$ pour N entier convenable (appendice A.2), telle que la multiplication et le passage à l'inverse soient de classe $\mathscr{C}^1$.

Un théorème fameux de Cartan assure que tout sous-groupe fermé de $\mathrm{GL}_n(\mathbb{K})$ est un groupe de Lie, mais sa preuve, très générale, n'est pas très éclairante et il est plus sportif de montrer sans ce théorème que les sous-groupes étudiés dans les exemples qui suivent vérifient bien les hypothèses. L'utilisation du théorème de submersion, théorème A.4, est alors essentielle.

1.2. Proposition. *Soit n un entier non nul. Le groupe $\mathrm{GL}_n(\mathbb{R})$ est une sous-variété de $\mathbb{R}^{n^2} \simeq \mathscr{M}_n(\mathbb{R})$. C'est un groupe de Lie.*

Démonstration. Comme $\mathrm{GL}_n(\mathbb{R})$ est un ouvert de l'espace $\mathscr{M}_n(\mathbb{R})$, il est clairement localement difféomorphe à $\mathbb{R}^{n^2}$; tous les ouverts d'une variété sont par définition des sous-variétés... De plus, le produit de matrices et le passage à l'inverse sont des applications polynomiales ou quotients de polynômes, donc de classe $\mathscr{C}^1$. □

1.3. Proposition. *L'espace tangent à $\mathrm{GL}_n(\mathbb{R})$ en l'identité $e = \mathrm{I}_n$ est :*

$$T_e\big(\mathrm{GL}_n(\mathbb{R})\big) = \mathscr{M}_n(\mathbb{R}).$$

Démonstration. Pour tout $X \in \mathscr{M}_n(\mathbb{R})$, il existe une courbe de classe $\mathscr{C}^1$, $\gamma : \mathbb{R} \to \mathrm{GL}_n(\mathbb{R})$, telle que $\gamma(0) = \mathrm{I}_n$ et $\gamma'(0) = X$. Par exemple, on peut prendre : $\gamma(t) = \exp(tX)$ $(t \in \mathbb{R})$. □

Notons que l'utilisation de l'exponentielle dans cette preuve n'est absolument pas nécessaire (comme $\mathrm{GL}_n(\mathbb{R})$ est un ouvert, on aurait pu aussi bien prendre un petit segment autour de I_n dirigé par X), mais elle est instructive.

Avant d'aller plus loin, faisons quelques échauffements avec des exemples de calculs de différentielles.

1.4. Exemples. On notera $\mathrm{d}f_A$ la différentielle de f en A.

(i) Soit $f : \mathscr{M}_n(\mathbb{R}) \to \mathscr{M}_n(\mathbb{R})$, $A \mapsto 2A$. Comme f est linéaire, elle est différentiable en tout point A et sa différentielle est $\mathrm{d}f_A : H \mapsto 2H$.

[1] C'est une définition *ad hoc* un peu trop restrictive.

(ii) Soit $f : \mathscr{M}_n(\mathbb{R}) \to \mathscr{M}_n(\mathbb{R})$, $A \mapsto A^2$. Pour $H \in \mathscr{M}_n(\mathbb{R})$, on calcule : $f(\mathrm{I}_n+H) = (\mathrm{I}_n+H)^2 = \mathrm{I}_n+2H+H^2$. La différentielle $\mathrm{d}f_{\mathrm{I}_n}$ correspond à la partie linéaire en H du développement, c'est : $\mathrm{d}f_{\mathrm{I}_n} : H \mapsto 2H$.

Pour obtenir $\mathrm{d}f_A$ pour A quelconque, ici et dans la suite, on prend la partie linéaire (en H) du développement de Taylor en A de $f(A+H)$. Cela a un sens par unicité du développement de Taylor.

Ici : $(A+H)^2 = A^2 + AH + HA + H^2$ pour tout (A, H), ce qui nous donne : $\mathrm{d}f_A : H \mapsto AH + HA$.

(iii) Soit $f : \mathrm{GL}_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R})$, $A \mapsto A^{-1}$. Pour H assez proche de 0, on a : $(\mathrm{I}_n + H)^{-1} = \mathrm{I}_n - H + H^2 - H^3 + \ldots$, et donc la différentielle de f est : $\mathrm{d}f_{\mathrm{I}_n} : H \mapsto -H$.

Pour $A \in \mathrm{GL}_n(\mathbb{R})$ quelconque, on calcule :

$$\begin{aligned}(A+H)^{-1} = \big[A(\mathrm{I}_n + A^{-1}H)\big]^{-1} &= (\mathrm{I}_n + A^{-1}H)^{-1}A^{-1} \\ &= A^{-1} - A^{-1}HA^{-1} + (A^{-1}H)^2A^{-1} + \ldots,\end{aligned}$$

si bien que la différentielle en A est[2] : $\mathrm{d}f_A : H \mapsto -A^{-1}HA^{-1}$.

(iv) Soit $f : \mathscr{M}_n(\mathbb{R}) \to \mathrm{End}\big(\mathscr{M}_n(\mathbb{R})\big)$, $A \mapsto \big(X \mapsto AX\big)$. Comme f est linéaire, sa différentielle en $A \in \mathscr{M}_n(\mathbb{R})$ est : $\mathrm{d}f_A = f$.

L'exemple qui suit sera particulièrement important[3] dans la suite.

(v) Soit

$$\begin{aligned} f : \mathrm{GL}_n(\mathbb{R}) &\longrightarrow \mathrm{Aut}\big(\mathscr{M}_n(\mathbb{R})\big) \\ A &\longmapsto \big(f_A : X \longrightarrow AXA^{-1}\big).\end{aligned}$$

Cette fois, f n'est pas linéaire. Si H est dans un voisinage convenable de 0, alors $\mathrm{I}_n + H$ est inversible. On calcule tout d'abord :

$$(\mathrm{I}_n + H)X(\mathrm{I}_n + H)^{-1} = \mathrm{I}_n + (HX - XH) + (-HXH + XH^2) + \ldots$$

L'identité de $\mathrm{GL}_n(\mathbb{R})$ est envoyée par f sur l'identité de $\mathrm{Aut}\big(\mathscr{M}_n(\mathbb{R})\big) \simeq \mathrm{GL}\big(\mathscr{M}_n(\mathbb{R})\big)$ et donc, par 1.3, $\mathrm{d}f_{\mathrm{I}_n}$ envoie l'espace tangent $\mathscr{M}_n(\mathbb{R})$ dans $\mathrm{End}\big(\mathscr{M}_n(\mathbb{R})\big)$. On doit maintenant décomposer l'automorphisme d'espace $X \mapsto (\mathrm{I}_n+H)X(\mathrm{I}_n+H)^{-1}$ en somme d'endomorphismes (en X). On écrit alors

$$f(\mathrm{I}_n + H) = f_{\mathrm{I}_n+H} = \mathrm{Id}_{\mathscr{M}_n(\mathbb{R})} + g_1 + g_2 + \ldots,$$

où $g_1(X) = HX - XH$, $g_2(X) = -HXH + XH^2$, les autres endomorphismes (en X) étant de degré supérieur en H. La différentielle de f

[2] Si $n = 1$, on retrouve le fait que la dérivée de $f : a \mapsto a^{-1}$ sur $\mathbb{R}^*$ est $f' : a \mapsto -a^{-2}$.

[3] En fait, toute la pertinence des algèbres de Lie repose dessus.

en l'identité est donc :

$$\begin{aligned} \mathrm{d}f_{\mathrm{I}_n} : \mathscr{M}_n(\mathbb{R}) &\longrightarrow \mathrm{End}\big(\mathscr{M}_n(\mathbb{R})\big) \\ H &\longmapsto HX - XH. \end{aligned}$$

1.5. Proposition. *Le groupe* $\mathrm{O}(n,\mathbb{R})$ *est une sous-variété de* $\mathbb{R}^{n^2}$ *de dimension* $n(n-1)/2$. *Son espace tangent en l'identité est le sous-espace* $\mathscr{A}_n(\mathbb{R})$ *des matrices antisymétriques.*

Démonstration. Considérons l'application

$$\begin{aligned} f : \mathscr{M}_n(\mathbb{R}) &\longrightarrow \mathscr{S}_n(\mathbb{R}) \\ M &\longmapsto {}^tMM. \end{aligned}$$

On sait que $\mathrm{O}(n,\mathbb{R}) = f^{-1}(\mathrm{I}_n)$. Soit $A \in \mathscr{M}_n(\mathbb{R})$ et $H \in \mathscr{M}_n(\mathbb{R})$. On a :

$$f(A+H) = {}^t(A+H)(A+H) = {}^tAA + {}^tAH + {}^tHA + H^2,$$

d'où la différentielle de f en A :

$$\begin{aligned} \mathrm{d}f_A : \mathscr{M}_n(\mathbb{R}) &\longrightarrow \mathscr{S}_n(\mathbb{R}) \\ H &\longmapsto {}^tAH + {}^tHA. \end{aligned}$$

Soit $S \in \mathscr{S}_n(\mathbb{R})$, $A \in \mathrm{O}(n,\mathbb{R})$ et posons $H = AS/2$. Alors,

$${}^tAH + {}^tHA = \frac{1}{2}\,{}^tAAS + \frac{1}{2}S\,{}^tAA = \frac{1}{2}S + \frac{1}{2}S = S.$$

Donc, $\mathrm{d}f_A$ est surjective pour $A \in \mathrm{O}(n,\mathbb{R})$. D'après le théorème de submersion A.4, le groupe $\mathrm{O}(n,\mathbb{R})$ est une sous-variété de $\mathscr{M}_n(\mathbb{R})$ de dimension :

$$\dim \mathscr{M}_n(\mathbb{R}) - \dim \mathscr{S}_n(\mathbb{R}) = n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2};$$

de plus :

$$\begin{aligned} T_{\mathrm{I}_n}\big(\mathrm{O}(n,\mathbb{R})\big) &= \mathrm{Ker}(\mathrm{d}f_{\mathrm{I}_n}) \\ &= \{H \in \mathscr{M}_n(\mathbb{R}) \,:\, H + {}^tH = 0\} \\ &= \mathscr{A}_n(\mathbb{R}). \end{aligned}$$

□

On notera par la lettre gothique $\mathfrak{g}$ correspondante l'espace tangent $T_e(G)$ à un groupe G en l'identité e. Par exemple, on note $\mathfrak{o}_n(\mathbb{R}) := T_{\mathrm{I}_n}\big(\mathrm{O}(n,\mathbb{R})\big)$.

1.6. Exercice. *Soit* p, q *deux entiers, et* $n = p+q$. *Montrer que* $\mathrm{O}(p,q)$ *est un groupe de Lie et que l'on a :*

$$\mathfrak{o}(p,q) = \big\{H \in \mathscr{M}_n(\mathbb{R}) \,:\, {}^tH\mathrm{I}_{(p,q)} + \mathrm{I}_{(p,q)}H = 0\big\}.$$

1.7. Proposition. *La différentielle du déterminant en $A \in \mathscr{M}_n(\mathbb{R})$ est $H \mapsto \operatorname{tr}(\tilde{A}H)$, où $\tilde{A}$ est la transposée de la comatrice de A.*

En particulier, $\mathrm{SL}_n(\mathbb{R})$ est une sous-variété de $\mathbb{R}^{n^2}$ et son espace tangent en l'identité est :

$$\mathfrak{sl}_n(\mathbb{R}) = \{M \in \mathscr{M}_n(\mathbb{R}),\ \operatorname{tr}(M) = 0\}.$$

Démonstration. Comme $f = \det$ est polynomiale, elle a une différentielle en tout point A : $\mathrm{d}f_A : \mathscr{M}_n(\mathbb{R}) \to \mathbb{R}$. Soit $H \in \mathscr{M}_n(\mathbb{R})$. Notons $(A_i)_{1 \leqslant i \leqslant n}$ les colonnes de A et $(H_i)_{1 \leqslant i \leqslant n}$ les colonnes de H, et le déterminant comme un produit vectoriel : $\det A = A_1 \wedge A_2 \wedge \cdots \wedge A_n$, alors :

$$f(A+H) = \det(A+H) = (A_1 + H_1) \wedge (A_2 + H_2) \wedge \cdots \wedge (A_n + H_n).$$

Par multilinéarité, on obtient en ne conservant que la partie linéaire en H :

$$\mathrm{d}f_A(H) = H_1 \wedge A_2 \wedge \cdots \wedge A_n + A_1 \wedge H_2 \wedge A_3 \wedge \cdots \wedge A_n + \cdots + A_1 \wedge \cdots \wedge A_{n-1} \wedge H_n.$$

Écrivons $(m_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}} = {}^t\mathrm{com}(A) = \tilde{A}$ et $H = (h_{ij})$. En développant le déterminant $A_1 \wedge \cdots \wedge H_i \wedge \cdots \wedge A_n$ par rapport à la i^{e} colonne, il vient :

$$\mathrm{d}f_A(H) = \sum_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}} h_{ij} m_{ji} = \operatorname{tr}(\tilde{A}H).$$

À présent, supposons que $A \in \mathrm{SL}_n(\mathbb{R})$. Par les formules de Cramer, on a : $\tilde{A} = A^{-1} \neq 0$. Par suite, la forme linéaire $\mathrm{d}f_A$ n'est pas nulle et donc, elle est surjective (l'espace d'arrivée est de dimension 1). Comme $\mathrm{SL}_n(\mathbb{R}) = f^{-1}(1)$, c'est, par le théorème de submersion A.4, une sous-variété de $\mathbb{R}^{n^2}$ de dimension $n^2 - 1$ et, de plus :

$$\mathfrak{sl}_n(\mathbb{R}) = \{H \in \mathscr{M}_n(\mathbb{R}) \ :\ \operatorname{tr}(H) = 0\}. \qquad \square$$

Démontrons rapidement des résultats analogues sur les complexes. Pour avoir une structure différentielle, on identifie $\mathscr{M}_n(\mathbb{C})$ à $\mathbb{R}^{2n^2}$.

1.8. Proposition. *Soit n, p, q des entiers naturels. Les groupes $\mathrm{GL}_n(\mathbb{C})$, $\mathrm{SL}_n(\mathbb{C})$, $\mathrm{O}_n(\mathbb{C})$, $\mathrm{SO}_n(\mathbb{C})$, $\mathrm{U}(n) = \mathrm{U}_n(\mathbb{C})$, $\mathrm{SU}(n) = \mathrm{SU}_n(\mathbb{C})$, $\mathrm{SU}(p,q)$ sont des groupes de Lie.*

Démonstration. Le produit de matrices et le passage à l'inverse sont différentiables sur $\mathbb{C}$ comme sur $\mathbb{R}$. Il suffit donc de montrer que les groupes sont des sous-variétés.

D'abord, le groupe linéaire $\mathrm{GL}_n(\mathbb{C})$ est un ouvert de $\mathscr{M}_n(\mathbb{C})$, donc c'est une sous-variété. Le groupe spécial linéaire complexe se traite comme le réel :

l'application $f = \det : \mathscr{M}_n(\mathbb{C}) \to \mathbb{C}$ est différentiable et sa différentielle en la matrice $A \in \mathrm{SL}_n(\mathbb{C})$ est la forme linéaire $H \mapsto \mathrm{tr}(A^{-1}H)$, qui est manifestement non nulle, et f est donc une submersion en tout point de $\mathrm{SL}_n(\mathbb{C})$. Par suite, $\mathrm{SL}_n(\mathbb{C})$ est une sous-variété de $\mathscr{M}_n(\mathbb{C})$.

Idem. Pour $\mathrm{O}_n(\mathbb{C})$, on peut juste relire la preuve réelle. Pour $\mathrm{U}_n(\mathbb{C})$, on l'adapte en introduisant $f : \mathscr{M}_n(\mathbb{C}) \to \mathscr{H}_n(\mathbb{C})$, $A \mapsto A^*A$. La différentielle de f en A est : $H \mapsto A^*H + H^*A$: si A est inversible, c'est la composée des surjections $\mathscr{M}_n(\mathbb{C}) \to \mathscr{M}_n(\mathbb{C})$, $H \mapsto A^*H$ et $\mathscr{M}_n(\mathbb{C}) \to \mathscr{H}_n(\mathbb{C})$, $K \mapsto K + K^*$. Ainsi, f est une submersion en tout point de $\mathrm{U}_n(\mathbb{C}) = f^{-1}(\mathrm{I}_n)$.

Le cas de $\mathrm{SU}_n(\mathbb{C})$, un peu plus délicat, sera traité en A.5. Le cas de $\mathrm{SU}(p,q)$ est laissé à titre d'exercice. □

1.9. Exercice. *Vérifier que pour n, p, q entiers, on a :*

$$\mathfrak{o}_n(\mathbb{C}) = \{H \in \mathscr{M}_n(\mathbb{R}) \,:\, {}^tH + H = 0\} = \mathscr{S}_n(\mathbb{C}),$$
$$\mathfrak{u}_n(\mathbb{C}) = \{H \in \mathscr{M}_n(\mathbb{R}) \,:\, H^* + H = 0\} = \mathscr{AH}_n,$$
$$\mathfrak{su}(p,q) = \{H \in \mathscr{M}_n(\mathbb{R}) \,:\, H^* + H = 0,\ \mathrm{tr}(H) = 0\},$$

où $\mathscr{AH}_n$ est le $\mathbb{R}$-espace vectoriel des matrices antihermitiennes.

1.10. Mise en garde. Attention toutefois : le théorème de submersion ne suffit pas toujours pour montrer qu'une intersection de sous-variétés est une sous-variété. On rencontre cette difficulté avec $\mathrm{SU}_n(\mathbb{C})$, car l'application $M \mapsto (M^*M, \det M)$ n'est pas une submersion vers $\mathscr{H}_n \times \mathbb{C}$. On pourrait utiliser une version plus générale du théorème de submersion, pas beaucoup plus difficile, où l'application de départ est une application entre sous-variétés. Ici, le déterminant est une submersion de $\mathrm{U}_n(\mathbb{C})$ vers le cercle $\mathbf{S}^1$. Mais cela sort du cadre des outils mis à disposition dans le présent livre. Voir le § A.5.

2. Applications aux isomorphismes exceptionnels

On va voir que le formalisme des groupes de Lie est efficace pour montrer la surjectivité dans les isomorphismes qui suivent. Cette proposition en est juste un exemple au même titre que la proposition qui suit. La méthode utilisée pour la surjectivité, et qui justifie l'utilisation des groupes de Lie, sera généralisée en appendice, voir corollaire IX-A.8.2.

2.1. Proposition. *L'action de* $\mathrm{SL}_2(\mathbb{C})$ *sur (son algèbre de Lie)* $\mathfrak{sl}_2(\mathbb{C})$ *par conjugaison fournit un isomorphisme :* $\mathrm{PSL}_2(\mathbb{C}) \simeq \mathrm{SO}_3(\mathbb{C})$.

Démonstration. Le groupe $\mathrm{SL}_2(\mathbb{C})$ agit sur $\mathfrak{sl}_2(\mathbb{C})$ par conjugaison :

$$\begin{aligned} \varphi : \mathrm{SL}_2(\mathbb{C}) &\longrightarrow \mathrm{Aut}\big(\mathfrak{sl}_2(\mathbb{C})\big) \\ A &\longmapsto (\varphi_A : X \longmapsto AXA^{-1}). \end{aligned}$$

C'est bien une action, car l'on a : $\mathrm{tr}(AXA^{-1}) = \mathrm{tr}(X) = 0$. On a de plus : $\dim \mathfrak{sl}_2(\mathbb{C}) = 3$, car $\mathfrak{sl}_2(\mathbb{C})$ est un hyperplan de $\mathscr{M}_2(\mathbb{C})$.

Puisque, pour tout A, l'application φ_A est linéaire et bien sûr bijective, on voit que $\varphi\big(\mathrm{SL}_2(\mathbb{C})\big)$ est contenu dans le groupe linéaire de l'espace vectoriel $\mathfrak{sl}_2(\mathbb{C})$, que l'on identifie à $\mathrm{GL}_3(\mathbb{C})$.

Plus précisément, puisque $\det\big(\varphi_A(X)\big) = \det(X)$, on peut même dire que $\varphi\big(\mathrm{SL}_2(\mathbb{C})\big)$ est un sous-groupe de $\mathrm{O}(\det)$, le groupe orthogonal de la forme déterminant (en dimension 2, c'est bien une forme quadratique).

Soit $X \in \mathfrak{sl}_2(\mathbb{C})$. Alors, on peut l'écrire

$$X = \begin{pmatrix} a & c \\ b & -a \end{pmatrix} \quad \text{où } a, b, c \in \mathbb{C}.$$

On calcule

$$\det(X) = -a^2 - bc = -a^2 - \frac{1}{4}(b+c)^2 + \frac{1}{4}(b-c)^2.$$

Les formes obtenues sont linéairement indépendantes, et la forme quadratique est donc non dégénérée, et de rang (maximal) 3. Dans une base orthonormée[4] de $\mathfrak{sl}_2(\mathbb{C})$ pour la forme quadratique det, le groupe orthogonal $\mathrm{O}(\det)$ s'identifie au groupe $\mathrm{O}_3(\mathbb{C})$ des matrices orthogonales complexes, si bien que l'on peut écrire $\varphi\big(\mathrm{SL}_2(\mathbb{C})\big) \subset \mathrm{O}_3(\mathbb{C})$. Puisque φ est continue et $\mathrm{SL}_2(\mathbb{C})$ connexe, $\varphi\big(\mathrm{SL}_2(\mathbb{C})\big)$ est aussi connexe dans $\mathrm{O}_3(\mathbb{C})$ tout en contenant l'identité, et donc $\varphi\big(\mathrm{SL}_2(\mathbb{C})\big) \subset \mathrm{SO}_3(\mathbb{C})$.

Montrons l'inclusion inverse. Composée de fractions rationnelles, φ est différentiable :

$$\begin{aligned} \mathrm{d}\varphi_{\mathrm{I}} : \mathfrak{sl}_2(\mathbb{C}) &\longrightarrow \mathfrak{so}_3(\mathbb{C}) \\ H &\longmapsto \mathrm{d}\varphi_{\mathrm{I}}(H) : X \longmapsto HX - XH. \end{aligned}$$

On a : $\mathfrak{so}_3(\mathbb{C}) = \mathfrak{o}_3(\mathbb{C})$ (exercice : utiliser le fait que $\mathrm{SO}_3(\mathbb{C})$ est la composante neutre dans $\mathrm{O}_3(\mathbb{C})$). De plus, $\mathrm{d}\varphi_{\mathrm{I}}$ est injective, car l'on a :

$$\begin{aligned} \mathrm{Ker}\,\mathrm{d}\varphi_{\mathrm{I}} &= \{H \in \mathfrak{sl}_2(\mathbb{C}) : \forall X,\ HX - XH = 0\} \\ &= \{H : \mathrm{tr}(H) = 0, H = \lambda \mathrm{I}_2\} = \{0\}. \end{aligned}$$

[4] Une telle base existe puisque l'on est sur $\mathbb{C}$. Souvenons-nous au passage que sur le corps $\mathbb{C}$ les formes quadratiques sont classées par le rang.

Comme $\dim \mathfrak{sl}_2(\mathbb{C}) = \dim \mathfrak{so}_3(\mathbb{C}) = 3$, on conclut que $\mathrm{d}\varphi_I$ est un isomorphisme. Ainsi, d'après le théorème d'inversion locale A.1, l'application φ est un homéomorphisme local et donc $\varphi\big(\mathrm{SL}_2(\mathbb{C})\big)$ contient un ouvert. Le sous-groupe $\varphi\big(\mathrm{SL}_2(\mathbb{C})\big)$ est ouvert par principe de translation, et ainsi nécessairement fermé également.

À la fois ouvert et fermé non vide dans $\mathrm{SO}_3(\mathbb{C})$, qui est connexe, on a donc finalement : $\varphi\big(\mathrm{SL}_2(\mathbb{C})\big) = \mathrm{SO}_3(\mathbb{C})$. Puisque $\operatorname{Ker}\varphi = \mathbb{K}^*\mathrm{I}_2$ (les homothéties non nulles), on peut conclure que $\mathrm{PSL}_2(\mathbb{C}) \simeq \mathrm{SO}_3(\mathbb{C})$. □

Ce type d'isomorphismes exceptionnels prend parfois le nom de « correspondance de Klein ». Suivent d'autres exemples en exercice. On pourra trouver l'équivalent sur les corps finis dans le second tome, chapitre XII.

2.2. Exercice. *Montrer que* $\mathrm{PSL}_2(\mathbb{R}) \simeq \mathrm{SO}_0(2,1)$ *par l'action de* $\mathrm{SL}_2(\mathbb{R})$ *sur* $\mathfrak{sl}_2(\mathbb{R})$ *par conjugaison.*

Montrer que $\mathrm{PSU}_2(\mathbb{C}) \simeq \mathrm{SO}_3(\mathbb{R})$ *par l'action de* $\mathrm{SU}_2(\mathbb{C})$ *sur l'espace (réel) des matrices hermitiennes de trace nulle* $\mathscr{H}_2(\mathbb{C}) \cap \mathfrak{sl}_2(\mathbb{C}) = \{H \in \mathscr{M}_2(\mathbb{C}) : H^* = H,\ \mathrm{tr}(H) = 0\}$ *par congruence hermitienne.*

2.3. Proposition. *L'action de* $\mathrm{SL}_2(\mathbb{C})$ *sur l'espace réel des matrices hermitiennes de taille* 2 *par congruence fournit un isomorphisme*

$$\mathrm{PSL}_2(\mathbb{C}) \simeq \mathrm{SO}_0(3,1).$$

Démonstration. Identique à la précédente démonstration. Le groupe $\mathrm{SL}_2(\mathbb{C})$ agit sur l'espace réel des matrices hermitiennes $\mathscr{H}_2(\mathbb{C})$ (de dimension 4 sur $\mathbb{R}$) par congruence hermitienne :

$$\begin{aligned} \varphi : \mathrm{SL}_2(\mathbb{C}) &\longrightarrow \mathrm{End}\big(\mathscr{H}_2(\mathbb{C})\big) \\ A &\longmapsto (\varphi_A : X \longmapsto AXA^*). \end{aligned}$$

L'image de φ est alors isomorphe à $\mathrm{SO}_0(q)$, où q est une forme quadratique sur $\mathscr{H}_2(\mathbb{C})$ de signature $(3,1)$.

Pour faire cette preuve soi-même, voici une méthode en quelques étapes.

1. Vérifier que c'est bien une action.
2. Montrer que φ_A respecte le déterminant pour tout A dans $\mathrm{SL}_2(\mathbb{C})$: $\det(AXA^*) = \det(X)$ pour tout X, car le déterminant de A vaut 1.
3. Calculer la signature de cette forme quadratique, c'est-à-dire celle de

$$\begin{vmatrix} a & b-ic \\ b+ic & d \end{vmatrix} = ad - (b^2 + c^2).$$

On doit trouver $(1,3)$. Bien sûr, si l'on prend $-\det$ on trouve $(3,1)$, ce qui revient au même (les deux formes ont le même groupe d'isométries).

4. Montrer que $\varphi\big(\mathrm{SL}_2(\mathbb{C})\big) \subset \mathrm{SO}_0(3,1)$ par un argument topologique.
5. Montrer que φ est différentiable et que sa différentielle en l'identité est :

$$\begin{aligned} \mathrm{d}\varphi_{\mathrm{I}_2} : \mathfrak{sl}_2(\mathbb{C}) &\longrightarrow \mathfrak{so}_0(3,1) \\ H &\longmapsto \mathrm{d}\varphi_{\mathrm{I}_2}(H), \quad \text{où} \quad \mathrm{d}\varphi_{\mathrm{I}_2}(H) : X \longmapsto HX + XH^*. \end{aligned}$$

6. Montrer qu'elle est injective ($\operatorname{Ker} \mathrm{d}\varphi_{\mathrm{I}_2} = 0$).

   On peut pour cela résoudre brutalement un système linéaire, car l'on est en basse dimension. On peut aussi appliquer le proposition V-B.2.2, *l'adaptation du cas symétrique au cas hermitien étant indolore.*
7. Montrer qu'elle est surjective par un argument de dimension (vérifier que l'on a : $\dim \mathfrak{sl}_2(\mathbb{C}) = \dim \mathfrak{so}_0(3,1)$).
8. Enfin, conclure en utilisant le théorème d'inversion locale et les propriétés des groupes topologiques. □

2.4. Remarque. On peut restreindre cette action à $\mathrm{SL}_2(\mathbb{R})$ pour retrouver le résultat de l'exercice 2.2, en observant que $\mathrm{SL}_2(\mathbb{R})$ stabilise l'hyperplan $\mathscr{S}_2(\mathbb{R})$, formé des matrices réelles symétriques, de $\mathscr{H}_2(\mathbb{C})$. Nous verrons dans le second volume comment relier cette action à la géométrie du demi-plan de Poincaré.

A. Annexe. Rappels et compléments en calcul différentiel

Dans cette partie, on suit le livre *Catastrophes et bifurcations* [12] de Michel Demazure.

Ici, $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

Si f est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$, on sait comment faire, à l'aide des théorèmes de base de l'algèbre linéaire, pour montrer que f est injective, surjective, ou pour calculer son rang...

Le calcul différentiel permet de généraliser ces caractérisations au cas où f est une application de classe $\mathscr{C}^1$ de $\mathbb{R}^n$ dans $\mathbb{R}^m$ (on note $f \in \mathscr{C}^1(\mathbb{R}^n, \mathbb{R}^m)$ munis de leur topologie usuelle). On utilise un principe fécond de *linéarisation.*

Une différence importante avec l'algèbre linéaire est que l'on obtient des résultats locaux. Le théorème qui règne en maître est le théorème d'inversion locale, dont les hypothèses généralisent celles d'un système de Cramer.

A.1. Théorème (d'inversion locale). *Soit $f \in \mathscr{C}^1(\mathbb{R}^n, \mathbb{R}^n)$ et $a \in \mathbb{R}^n$. Si la différentielle $\mathrm{d}f_a$ de f en a est un isomorphisme linéaire, alors il existe un voisinage V de a tel que la restriction de f à V soit un difféomorphisme de V vers un voisinage ouvert de $f(a)$.*

Pour appliquer ce résultat puissant aux groupes de Lie classiques, le bon cadre va être de travailler sur des variétés différentiables, ou plutôt dans le cadre un peu plus simple des sous-variétés de $\mathbb{R}^n$.

A.2. Définitions et premières propriétés

A.2.1. Définition. Soit $n \in \mathbb{N}^*$. On dit qu'une partie M de $\mathbb{R}^n$ est une *sous-variété* si elle est une sous-variété en chacun de ses points. On dit que M est une sous-variété de dimension m de $\mathbb{R}^n$ en l'un de ses points x_0 s'il existe un voisinage W de x_0 et un difféomorphisme $f : W \to U$ sur un ouvert U de $\mathbb{R}^n$ qui envoie x_0 sur 0, tel que $M \cap W = f^{-1}(\{0\} \times \mathbb{R}^m)$ (ici, on identifie $\mathbb{R}^n$ avec $\mathbb{R}^{n-m} \times \mathbb{R}^m$).

Le caractère « localement $\mathbb{R}^m$ » de la définition d'une sous-variété va faire que le théorème d'inversion locale n'a aucun mal à s'y généraliser. Maintenant, cette définition peut en rebuter quelques-uns, mais pour rassurer les uns et informer les autres : d'une part, elle est assez naturelle (si, si !), puisqu'une structure de variété sur X n'est rien d'autre qu'un atlas avec changement de cartes de classe $\mathscr{C}^1$; une structure de sous-variété de $\mathbb{R}^n$ sur X est donc donnée par un atlas, mais cette structure doit provenir d'un atlas sur $\mathbb{R}^n$. D'autre part, on ne va pas s'en servir concrètement grâce au théorème de submersion, qui est un corollaire du théorème d'inversion locale.

La définition de la sous-variété M donne une présentation par un système d'équations au voisinage de x_0 : on note les composantes de f ainsi :

$$\forall x \in W, \quad f(x) = \big(f_1(x), \dots, f_n(x)\big),$$

alors $M \cap W$ est l'ensemble des x qui sont les solutions de

$$(*) \qquad f_1(x) = f_2(x) = \cdots = f_{n-m}(x) = 0.$$

Remarquons que le système $(*)$ est non dégénéré, au sens où les différentielles des f_i en x_0 sont linéairement indépendantes : par exemple, on peut dire que leurs matrices sont les lignes de la matrice jacobienne de f en x_0, qui est inversible.

A.2.2. Proposition. *Soit $f_1, \dots, f_{n-m}$ des fonctions différentiables définies sur un voisinage W d'un point x_0 de $\mathbb{R}^d$. On suppose d'une part que $f_1(x_0) = \cdots = f_{d-m}(x_0) = 0$ (c'est accessoire) et d'autre part que les différentielles $(\mathrm{d}f_i)_{x_0}$ $(1 \leqslant i \leqslant n-m)$ sont linéairement indépendantes. Alors, l'ensemble M des solutions du système $(*)$ ci-dessus est une sous-variété en x_0.*

Démonstration. En effet, par le théorème de la base incomplète, il existe des formes linéaires $u_1, \dots, u_m$ sur $\mathbb{R}^n$ telles que

$$\big((\mathrm{d}f_1)_{x_0}, \dots, (\mathrm{d}f_{n-m})_{x_0}, u_1, \dots, u_m\big)$$

soit une base de l'espace dual de $\mathbb{R}^n$. Mais alors, les fonctions (linéaires) $f_{n-m+1} = u_1, \dots, f_n = u_m$ sont différentiables et égales à leur différentielle (calculée en n'importe quel point) et la fonction $f = (f_1, \dots, f_n) : W \to \mathbb{R}^n$ est une application différentiable dont la différentielle en x_0 est inversible, donc un difféomorphisme local de W sur un voisinage de 0. □

A.3. Espace tangent

On travaille avec des sous-variétés. Il est, en effet, relativement facile de définir l'espace tangent en un point courant d'une sous-variété. Ces espaces tangents offrent alors une bonne « approximation » de la sous-variété et permettent donc d'appliquer ce principe essentiel qu'est la « linéarisation », principe qui remonte déjà à Newton et Leibniz.

Voici une première définition intuitive de l'espace tangent.

A.3.1. Définition. Soit X une sous-variété de $\mathbb{R}^n$ et $x \in X$. On appelle espace tangent en x à X, et l'on note $T_x(X)$, l'ensemble des vecteurs $\gamma'(0)$ tangents en x aux courbes $\gamma : \mathbb{R} \to \mathbb{R}^n$ de classe $\mathscr{C}^1$ passant par x et tracées dans X (courbes que l'on peut voir donc comme des fonctions $\gamma \in \mathscr{C}^1(\mathbb{R}, \mathbb{R}^n)$ dont l'image est incluse dans X et telles que $\gamma(0) = x$).

On montre que cette définition est équivalente à celle-ci, plus subtile[5].

A.3.2. Définition. On reprend les notations de la définition A.2.1. L'espace tangent à M en x_0 est l'image de $\{0\} \times \mathbb{R}^m$ par l'inverse $(\mathrm{d}f_{x_0})^{-1}$ de la différentielle de f en x_0. C'est aussi l'intersection des noyaux des formes linéaires qui apparaissent dans l'expression ci-après :

$$T_{x_0}M = \bigcap_{i=1}^{d-m} \mathrm{Ker}(\mathrm{d}f_i)_{x_0}.$$

[5] Elle n'est pas si mystérieuse pour autant : l'espace tangent en l'origine à $\{0\} \times \mathbb{R}^m$ est $\{0\} \times \mathbb{R}^m$ lui-même ; on le transporte par la différentielle du difféomorphisme f.

Remarquons que cet espace est également l'intersection des noyaux des différentielles en x_0 de toutes les fonctions différentiables nulles sur M.

On vérifie que l'espace tangent est bien un sous-espace vectoriel dont la dimension est celle de la variété. Le théorème d'inversion locale s'exprime, dans le cadre des sous-variétés de $\mathbb{R}^n$, de la façon suivante.

A.3.3. Théorème (d'inversion locale, *bis*). *Soit X et Y des sous-variétés de $\mathbb{R}^n$ et $\mathbb{R}^m$, respectivement. Soit $f \in \mathscr{C}^1(X,Y)$ et $a \in X$.*

Si $\mathrm{d}f_a$ est un isomorphisme de $T_a(X)$ sur $T_{f(a)}(Y)$, alors il existe un voisinage $V \subset X$ de a tel que la restriction de f à V soit un difféomorphisme de V vers un ouvert $W \subset Y$ de $f(a)$.

A.4. Théorème de submersion

A.4.1. Définition. Soit $f \in \mathscr{C}^1(\mathbb{R}^n, \mathbb{R}^m)$ et W un ouvert de $\mathbb{R}^n$. On dit que f est une submersion en un point x_0 de W si la différentielle de f en x_0 est une surjection de $\mathbb{R}^n$ sur $\mathbb{R}^m$ (en particulier, $n \geqslant m$).

On dit que f est une submersion si f est une submersion en tout point.

A.4.2. Théorème (de submersion).

Soit $f \in \mathscr{C}^1(\mathbb{R}^n, \mathbb{R}^m)$ et $X := f^{-1}(\{y\})$, où $y \in \mathbb{R}^m$.

*Si f est une submersion en tout point de X, (*i.e. *si $\mathrm{d}f_x$ est surjective en tout point x de X), alors X est une sous-variété de $\mathbb{R}^n$ de dimension $n-m$.*

A.4.3. Mise en garde. On n'a pas l'équivalent en immersion, c'est-à-dire que si $\mathrm{d}f_x$ est injective en tout point de $\mathbb{R}^n$, l'image de f n'est pas nécessairement une sous-variété (penser à une lemniscate, et méditer).

Démonstration. Ce théorème provient d'une propriété plus générale.

Soit W un ouvert de $\mathbb{R}^n$ contenant x_0 et $f : W \to \mathbb{R}^m$ ($m \in \mathbb{N}^*$) une application différentiable et submersive en x_0. Alors, la ligne de niveau de f à laquelle appartient x_0,

$$M = f^{-1}\big(\{f(x_0)\}\big),$$

est une sous-variété en x_0 de dimension $n-m$ dont l'espace tangent en x_0 est :

$$T_{x_0}M = \operatorname{Ker} \mathrm{d}f_{x_0}.$$

En effet, si l'on note $f = (f_1, \ldots, f_m)$ les composantes de f, les matrices des $(\mathrm{d}f_i)_{x_0}$ sont les lignes de la matrice jacobienne en x_0, c'est-à-dire de l'application linéaire $\mathrm{d}f_{x_0}$ dans les bases canoniques respectives. Puisque, par hypothèse, la matrice jacobienne est de rang m, ses m lignes sont linéairement indépendantes. On conclut par A.2.2. □

A.5. Application au groupe spécial unitaire

On peut montrer que le groupe unitaire $\mathrm{U}_n(\mathbb{C})$ est une sous-variété de $\mathscr{M}_n(\mathbb{C})$ en constatant que c'est l'image réciproque de l'identité I_n par la submersion

$$f : \mathrm{GL}_n(\mathbb{C}) \to \mathscr{H}_n, \quad f(g) = g^* g,$$

où $\mathscr{H}_n$ désigne l'espace des matrices hermitiennes. Voici la différentielle de f en I_n : $\mathrm{d}f_{\mathrm{I}_n}(H) = H^* + H$. C'est une surjection puisqu'une matrice hermitienne A est l'image de $A/2$.

Pour $\mathrm{SU}_n(\mathbb{C}) = \mathrm{U}_n(\mathbb{C}) \cap \mathrm{SL}_n(\mathbb{C})$, on peut penser à la fonction $F = (f, \det) : \mathrm{GL}_n(\mathbb{C}) \to \mathscr{H}_n \times \mathbb{C}^*$. Malheureusement, cette fonction n'est pas submersive. En effet, sa différentielle en I_n est l'application

$$\mathrm{d}F_{\mathrm{I}_n} : \mathscr{M}_n \to \mathscr{H}_n \times \mathbb{C}, \quad H \mapsto (H + H^*, \operatorname{tr} H).$$

Or, la trace d'une matrice hermitienne est réelle, et les éléments $(A, i\lambda)$ avec $A \in \mathscr{H}_n$ et $\lambda \in \mathbb{R}$ n'ont donc pas d'antécédent par $\mathrm{d}F_{\mathrm{I}_n}$. On va s'en tirer, au prix de renoncer à un traitement global.

A.5.1. Lemme. *Soit n un entier naturel non nul. Alors, le groupe $\mathrm{SU}_n(\mathbb{C})$ est une sous-variété de $\mathrm{GL}_n(\mathbb{C})$.*

Démonstration. Comme le déterminant d'une matrice unitaire a pour module 1, il suffit, pour assurer qu'elle appartient à $\mathrm{SU}_n(\mathbb{C})$, de vérifier que l'argument du déterminant est nul. Mais l'argument n'est pas défini sur $\mathbb{C}^*$, d'où le besoin de se restreindre à un voisinage de I_n.

On montre donc que $\mathrm{SU}_n(\mathbb{C})$ est une sous-variété en I_n. Puisque $\det \mathrm{I}_n = 1$, il existe un voisinage W de I_n sur lequel le déterminant ne prend aucune valeur dans $\mathbb{R}^-$. Notons $\arg : \mathbb{C} \setminus \mathbb{R}^- \to]-\pi, \pi[$ la détermination principale de l'argument. C'est la partie imaginaire de la réciproque de l'exponentielle $\exp :]0, +\infty[+ i\,]-\pi, \pi[\to \mathbb{C} \setminus \mathbb{R}^-$. Pour $z \in \mathbb{C} \setminus \mathbb{R}^-$, on a bien : $\arg z = \operatorname{Im} \exp^{-1}(z)$. Donc, arg est différentiable et sa différentielle en 1 est :

$$\forall h \in \mathbb{C}, \quad \mathrm{d}\arg_1(h) = (\mathrm{d}\operatorname{Im}\exp^{-1})_1(h) = (\mathrm{d}\operatorname{Im})_0[(\mathrm{d}\exp)_0]^{-1}(h) = \operatorname{Im}(h).$$

Par composition, il vient :

$$\forall H \in \mathscr{M}_n(\mathbb{C}), \quad \mathrm{d}(\arg \circ \det)_{\mathrm{I}_n}(H) = (\mathrm{d}\arg)_1 \circ (\mathrm{d}\det)_{\mathrm{I}_n}(H) = \operatorname{Im} \operatorname{tr} H.$$

Comme remarqué ci-dessus, $\mathrm{SU}_n(\mathbb{C}) \cap W$ est l'image réciproque du singleton $\{(\mathrm{I}_n, 0)\}$ par l'application :

$$G : W \to \mathscr{H}_n \times \mathbb{R}, \quad g \mapsto (g^* g, \arg\det g),$$

dont la différentielle en I_n est :

$$\mathrm{d}G_{\mathrm{I}_n}(H) = (H^* + H, \operatorname{Im}\operatorname{tr} H).$$

Pour $(A, \lambda) \in \mathscr{H}_n \times \mathbb{R}$, on a d'évidence :

$$A = B + B^*, \quad \lambda = \operatorname{Im}\operatorname{tr} B, \quad \text{où } B = \frac{1}{2}A + \frac{i\lambda}{n}\mathrm{I}_n.$$

Cela montre que G est une submersion en I_n, d'où $\mathrm{SU}_n(\mathbb{C})$ est une sous-variété en I_n.

À présent, soit $g \in \mathrm{SU}_n(\mathbb{C})$. La translation par g,

$$L_g : \mathrm{GL}_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C}), \quad h \mapsto gh$$

est un difféomorphisme et $\mathrm{SU}_n(\mathbb{C})$ est une sous-variété en I_n, donc l'image $L_g\big(\mathrm{SU}_n(\mathbb{C})\big) = \mathrm{SU}_n(\mathbb{C})$ est une sous-variété en $L_g(\mathrm{I}_n) = g$. □

A.6. Aparté : sous-groupes à un paramètre

Soit G un sous-groupe fermé de $\mathrm{GL}_n(\mathbb{K})$, où $\mathbb{K}$ est le corps réel ou complexe. On appelle sous-groupe à un paramètre de G un morphisme continu de $\mathbb{R}$ dans G.

A.6.1. Proposition. *Soit $\varphi : \mathbb{R} \to \mathrm{GL}_n(\mathbb{K})$ un morphisme continu. Il existe une matrice $X \in \mathscr{M}_n(\mathbb{K})$ telle que*

$$\forall t \in \mathbb{R}, \quad \varphi(t) = \exp(tX).$$

Démonstration. Montrons tout d'abord qu'un morphisme continu est automatiquement différentiable, ce qui en soi ne laisse pas de surprendre. On a, pour tout $(s,t) \in \mathbb{R}^2$:

$$\varphi(s+t) = \varphi(s)\varphi(t).$$

Par continuité de φ, on peut intégrer cette relation : pour $a > 0$, on a :

$$\forall s \in \mathbb{R}, \int_s^{s+a} \varphi(u)\,\mathrm{d}u = \int_0^a \varphi(s+t)\,\mathrm{d}t = \int_0^a \varphi(s)\varphi(t)\,\mathrm{d}t = \varphi(s)\int_0^a \varphi(t)\,\mathrm{d}t.$$

La fonction dans le membre de gauche est dérivable par rapport à s. Il suffit donc de montrer que $\int_0^a \varphi\,\mathrm{d}t$ est inversible pour a assez petit.

Fixons une norme sur l'espace des matrices. Il existe une boule centrée en I_n et contenue dans $\mathrm{GL}_n(\mathbb{K})$. Pour a assez petit et pour $t \in [0,a]$, $\varphi(t)$ appartient à cette boule ; *a fortiori*, la valeur moyenne de φ sur $[0,a]$,

$$\frac{1}{a}\int_0^a \varphi(t)\,\mathrm{d}t,$$

appartient à cette boule (convexité, si l'on veut), donc $\int_0^a \varphi$ est inversible.

Dérivons la relation de morphisme par rapport à s et prenons $s = 0$:

$$\forall t \in \mathbb{R}, \quad \varphi'(t) = \varphi'(0)\varphi(t).$$

On sait que l'unique solution de cette équation différentielle qui prend la valeur I_n en $t = 0$ est : $\varphi(t) = \exp tX$, où $X = \varphi'(0)$. □

A.7. Espace tangent en l'identité à un groupe de Lie

Dans ce texte, on appelle groupe de Lie un sous-groupe de $\mathrm{GL}_n(\mathbb{K})$ qui est aussi une sous-variété de $\mathscr{M}_n(\mathbb{K})$. Soit G un groupe de Lie dans $\mathrm{GL}_n(\mathbb{K})$.

A.7.1. Remarque. Le théorème de Cartan déjà cité prouve que tout sous-groupe fermé est automatiquement une sous-variété (une condition topologique implique une propriété différentielle, comme dans le paragraphe précédent). Cependant, on n'utilisera pas ce théorème : sa preuve est peu éclairante et il est assez facile de montrer au cas par cas que les groupes classiques sont des sous-variétés.

On note $e = \mathrm{I}_n$ le neutre de GL_n. Par définition, l'espace tangent à G en e est l'espace engendré par les vecteurs tangents en e aux courbes $\mathscr{C}^1$ tracées dans G et contenant e :

$$T_eG = \{\gamma'(0), \quad \gamma :]-\epsilon, \epsilon[\to G, \ \mathscr{C}^1, \ \gamma(0) = e\}.$$

A.7.2. Proposition. *Pour G un groupe de Lie dans $\mathrm{GL}_n(\mathbb{K})$ et $e = \mathrm{I}_n$, on a :*

$$T_eG = \{X \in \mathscr{M}_n(\mathbb{K}), \ \forall t \in \mathbb{R}, \ \exp(tX) \in G\}.$$

Démonstration. Soit $\mathfrak{g}$ l'ensemble des X tels que le sous-groupe à un paramètre $\{\exp tX\}$ soit contenu dans G. Comme $\gamma(t) = \exp tX$ est une courbe satisfaisant à $\gamma(0) = e$, il est clair que pour $X \in \mathfrak{g}$, on a : $X = \gamma'(0) \in T_eG$.

Inversement, soit $X \in T_eG$. On fixe une courbe $\gamma :]-\epsilon, \epsilon[\to G$ de classe $\mathscr{C}^1$, définie sur un voisinage de 0, telle que $\gamma(0) = e$ et $\gamma'(0) = X$.

On sait que l'exponentielle $\exp : \mathscr{M}_n(\mathbb{K}) \to \mathrm{GL}_n(\mathbb{K})$ a pour différentielle Id en 0, donc c'est un difféomorphisme d'un voisinage Ω de 0 dans $\mathscr{M}_n(\mathbb{K})$ vers un voisinage Ω' de e dans $\mathrm{GL}_n(\mathbb{K})$. Notons $\log : \Omega' \to \Omega$ le difféomorphisme réciproque. Pour m supérieur à m_0 convenable, $\gamma(1/m)$ appartient à Ω', ce qui donne un sens à $X_m = \log \gamma(1/m)$. Noter que $\exp X_m = \gamma(1/m) \in G$.

La suite $(X_m/||X_m||)_{m \geqslant m_0}$ prend ses valeurs dans la sphère unité, qui est compacte. Quitte à extraire une sous-suite, on peut donc supposer qu'elle converge.

Soit alors $t \in \mathbb{R}$. Pour $m \geqslant m_0$, écrivons

$$\frac{t}{||X_m||} = \lambda_m + \mu_m, \quad \text{où } \lambda_m \in \mathbb{Z} \text{ et } \mu_m \in [0,1[\,.$$

Comme la suite $(\exp X_m) = \big(\gamma(1/m)\big)$ converge vers e, la suite (X_m) tend vers 0. Comme (μ_m) bornée, la suite $(\exp \mu_m X_m)$ converge vers e. Et donc, les suites $\big(\exp(tX_m/||X_m||)\big)$ et $\big(\exp(\lambda_m X_m)\big)$ convergent vers la même limite.

Or, pour tout m, l'élément $\exp(\lambda_m X_m) = \exp(X_m)^{\lambda_m}$ est dans G, lequel est fermé. Par suite, la limite appartient aussi à G. Ainsi,

$$\lim_{m \to +\infty} \exp t \frac{X_m}{||X_m||} \in G.$$

D'autre part, pour m assez grand, on constate que

$$X_m = \frac{1}{m}\, \mathrm{d}\log \gamma'(0) + o\big(\frac{1}{m}\big) = \frac{1}{m}\, X + o\big(\frac{1}{m}\big).$$

On en déduit que la limite ci-dessus est :

$$\lim_{m \to +\infty} \exp t \frac{X_m}{||X_m||} = \exp t \frac{X}{||X||},$$

d'où l'on déduit que pour tout $t \in \mathbb{R}$, $\exp(tX/||X||)$ appartient à G, puis que X appartient à $\mathfrak{g}$. □

A.8. Noyau de la différentielle d'un morphisme

A.8.1. Proposition. *Soit $\varphi : G \to H$ un morphisme de classe $\mathscr{C}^1$ entre groupes de Lie. Soit e le neutre de G et $\mathrm{d}\varphi_e$ la différentielle de φ en e. Alors :*

$$\forall X \in \operatorname{Ker} \mathrm{d}\varphi_e,\ \forall t \in \mathbb{R}, \quad \exp tX \in \operatorname{Ker} \varphi.$$

Démonstration. La différentielle de φ en e est une application linéaire de T_eG dans T_eH. Soit $X \in \operatorname{Ker} \mathrm{d}\varphi_e$. D'après la caractérisation de l'espace tangent ci-dessus, on a : $\exp tX \in G$ pour tout t réel. On définit donc un sous-groupe à un paramètre par

$$\chi : \mathbb{R} \to H, \quad t \mapsto \varphi(\exp tX).$$

D'après la caractérisation des sous-groupes à un paramètre, on a :

$$\forall t \in \mathbb{R}, \quad \chi(t) = \exp t\chi'(0), \quad \text{où } \chi'(0) = \mathrm{d}\varphi_e(X).$$

À présent, si $X \in \operatorname{Ker} \mathrm{d}\varphi_e$, alors χ est constante, d'où $\exp(tX)$ appartient au noyau de φ. □

A.8.2. Corollaire. *Soit $\varphi : G \to H$ un morphisme de classe $\mathscr{C}^1$ entre groupes de Lie. Si le noyau de φ est discret, alors la différentielle de φ en le neutre e est injective. Si, de plus, les espaces tangents à G et H en leurs neutres ont la même dimension, alors φ est un difféomorphisme local. Si, de plus, H est connexe, alors φ est surjectif.*

Démonstration. Supposons que le noyau de φ est discret. Pour X dans le noyau de $\mathrm{d}\varphi_e$, la proposition montre que le sous-groupe à un paramètre $\chi : t \mapsto \exp(tX)$ prend ses valeurs dans $\operatorname{Ker}\varphi$. Comme le neutre est isolé dans le noyau, par connexité, χ est constant. Il en résulte que $X = \chi'(0) = 0$, ce qui prouve l'injectivité de la différentielle.

Lorsque les dimensions des espaces tangents au départ et à l'arrivée sont égales, la différentielle en e est nécessairement surjective. Par le théorème d'inversion locale, φ est un difféomorphisme local sur un voisinage du neutre de G.

Il en résulte que l'image de φ admet un intérieur non vide. Mais un tel sous-groupe est automatiquement ouvert (utiliser les translations !), si bien que l'image de φ est ouverte dans H. Elle est donc automatiquement fermée. En tenant compte de ce que H est connexe, il en résulte que φ est surjectif. □

B. Notion d'algèbre de Lie

Cette notion est hors-programme parce qu'il faut bien faire des choix et que l'esprit actuel de l'oral d'agrégation est prioritaire dans ce cours. Toutefois, nous engageons vivement la lecture de ce passage. Elle apporte un éclairage sur ce qui fait l'efficacité de la théorie des groupes de Lie : la linéarisation des problèmes sur le groupe G en considérant son espace tangent $\mathfrak{g}$ en l'identité, espace que l'on munit naturellement une structure d'algèbre non-associative, appelée structure d'algèbre de Lie. L'algèbre de Lie $\mathfrak{g}$ devient un recueil d'informations sur le groupe G.

B.1. Définition. Une *algèbre de Lie* est un espace vectoriel $\mathfrak{g}$ muni d'une application bilinéaire appelé *crochet de Lie* :

$$[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathfrak{g},\ (x,y) \longmapsto [x,y],$$

telle que pour tout (x,y,z) d'éléments de $\mathfrak{g}$, on ait :

$$\begin{cases} [x,y] = -[y,x] & \text{(antisymétrie)} \\ [[x,y],z] + [[z,x],y] + [[y,z],x] = 0 & \text{(identité de Jacobi).} \end{cases}$$

B.2. Remarque. Observer que l'identité de Jacobi est une somme cyclique sur les variables x, y et z.

B.3. Exemple. L'algèbre $\mathscr{M}_n(\mathbb{K})$, et plus généralement toute algèbre associative, est une algèbre de Lie pour le *commutateur* par $[A,B] = AB - BA$ pour deux éléments quelconques A et B.

B.4. Mise en garde. Attention à la structure considérée : le commutateur de deux éléments A et B dans une algèbre associative est : $[A,B] = AB - BA$, mais dans un groupe, c'est : $[A,B] = ABA^{-1}B^{-1}$.

Lorsque l'on manipule des matrices, où algèbres et groupes coexistent, cela peut prêter à confusion.

B.5. Proposition. *Soit G un sous-groupe de Lie de $\mathrm{GL}_n(\mathbb{K})$ (où le corps $\mathbb{K}$ est réel ou complexe). Alors,* $\mathrm{Lie}\, G = \mathfrak{g} := T_e(G) \subset \mathscr{M}_n(\mathbb{K})$ *possède une structure de sous-algèbre de Lie de $\mathscr{M}_n(\mathbb{K})$. Autrement dit :*

$$\forall X, Y \in \mathfrak{g}, \quad [X,Y] = XY - YX \in \mathfrak{g}.$$

On dit que $\mathfrak{g}$ est stable par le crochet $[\cdot,\cdot]$.

B.6. Définition. On dit que $(\mathfrak{g}, +, [\cdot,\cdot])$, ou, par abus de notation, $\mathfrak{g}$ est l'algèbre de Lie du groupe G.

Démonstration. Le groupe G agit sur lui-même par conjugaison en fixant la matrice I_n. Pour être plus explicite :

$$\begin{array}{rcl} \mathrm{Ad} : G & \longrightarrow & \mathrm{Aut}(G), \\ g & \longmapsto & \mathrm{Ad}_g \end{array} \quad \text{où} \quad \begin{array}{rcl} \mathrm{Ad}_g : G & \longrightarrow & \mathrm{Aut}(G) \\ h & \longmapsto & ghg^{-1}. \end{array}$$

En prenant la différentielle de Ad_g en l'identité, on obtient une application linéaire de l'espace tangent $\mathfrak{g}$ à G en l'identité dans lui-même. On obtient de plus une action linéaire de G sur $\mathfrak{g}$; notons-la φ :

$$\begin{array}{rcl} \varphi : G & \longrightarrow & \mathrm{GL}(\mathfrak{g}), \\ g & \longmapsto & \varphi(g) \end{array} \quad \text{où} \quad \begin{array}{rcl} \varphi(g) : \mathfrak{g} & \longrightarrow & \mathfrak{g} \\ X & \longmapsto & gXg^{-1}. \end{array}$$

On a $\varphi(\mathrm{I}_n) = \mathrm{I}_n$, donc

$$\begin{array}{rcl} \mathrm{d}\varphi_{\mathrm{I}_n} : \mathfrak{g} & \longrightarrow & \mathscr{M}_n(\mathfrak{g}) \\ X & \longmapsto & \mathrm{d}\varphi_{\mathrm{I}_n}(X). \end{array}$$

On a $\varphi(\mathrm{I}_n) = \mathrm{I}_n$, donc, en prenant la différentielle de Ad, on obtient :

$$\begin{aligned} \mathrm{d}\,\mathrm{Ad}_{\mathrm{I}_n} : \mathfrak{g} &\longrightarrow \mathscr{M}_n(\mathfrak{g}) \\ X &\longmapsto \mathrm{d}\varphi_{\mathrm{I}_n}(X). \end{aligned}$$

Cette différentielle est la partie linéaire en X de

$$\begin{aligned} Y \longmapsto (\mathrm{I}_n + X)Y(\mathrm{I}_n + X)^{-1} &= (Y + XY)(\mathrm{I}_n - X + X^2 + \ldots) \\ &= Y + XY - YX - XYX + YX^2 + \ldots \end{aligned}$$

On a donc, comme dans l'exemple 1.4, $\mathrm{d}\varphi_{\mathrm{I}_n}(X) : Y \longmapsto YX - XY$. Ainsi, $\mathfrak{g}$ est bien stable par le crochet, puisque $\mathrm{d}\varphi_{\mathrm{I}_n}(X)$ va de $\mathfrak{g}$ dans $\mathfrak{g}$. □

B.7. Exercice. *Vérifier directement cette proposition sur* $\mathfrak{sl}_n(\mathbb{K})$, $\mathfrak{so}_n(\mathbb{K})$, $\mathfrak{so}(p,q)$ *à l'aide des exemples calculés dans la section qui précède.*

On vient de rencontrer une nouvelle structure, celle d'algèbre de Lie, qui apparaît donc naturellement sur l'espace tangent au neutre e d'un groupe de Lie. Avec une nouvelle structure, arrive toujours la notion de morphisme.

B.8. Définition. Soit $\mathfrak{g}$ et $\mathfrak{h}$ deux algèbres de Lie. Un morphisme d'algèbres de Lie $\varphi : \mathfrak{g} \to \mathfrak{h}$ est un morphisme d'espaces vérifiant

$$[\varphi(x), \varphi(y)] = \varphi([x,y]), \quad x, y \in \mathfrak{g}.$$

La notion de sous-algèbre de Lie est naturelle : la partie $\mathfrak{h}$ est une sous-algèbre de l'algèbre de Lie $\mathfrak{g}$ si $\mathfrak{h}$ est un sous-espace vectoriel de $\mathfrak{g}$ stable par le crochet.

Pour définir la notion d'algèbre de Lie quotient, il faut d'abord définir celle d'idéal de Lie.

B.9. Définition. Une partie $\mathfrak{h}$ est un idéal de Lie de l'algèbre de Lie $\mathfrak{g}$ si $\mathfrak{h}$ est un sous-espace vectoriel de $\mathfrak{g}$ vérifiant

$$\forall x \in \mathfrak{g},\ \forall y \in \mathfrak{h}, \quad [x,y] \in \mathfrak{h}.$$

On montre facilement que si $\mathfrak{h}$ est un idéal de Lie de l'algèbre de Lie $\mathfrak{g}$, alors l'espace quotient $\mathfrak{g}/\mathfrak{h}$ admet une structure d'algèbre de Lie par le crochet de $\mathfrak{g}$ passé au quotient.

Soit φ un morphisme entre deux algèbres de Lie $\mathfrak{h}$ et $\mathfrak{g}$. On montre alors que $\operatorname{Ker}\varphi$ est un idéal de Lie de $\mathfrak{h}$, que $\operatorname{Im}(\varphi)$ est une sous-algèbre de Lie de $\mathfrak{g}$. Tout est prêt pour l'isomorphisme canonique d'algèbres de Lie

$$\mathfrak{h}/\operatorname{Ker}\varphi \simeq \operatorname{Im}(\varphi).$$

Nous allons maintenant admettre quelques résultats importants de la théorie. On recommande le chapitre 3 de [25] pour des preuves.

Cette proposition est somme toute bien naturelle.

B.10. Proposition. *La différentielle en l'identité d'un morphisme de groupes de Lie $\varphi : G \to H$ induit un morphisme entre leurs algèbres de Lie $\mathrm{d}\varphi : \mathfrak{g} \to \mathfrak{h}$.*

Ce théorème est fondamental pour faire « remonter » les informations de l'algèbre de Lie à son groupe de Lie.

B.11. Théorème. *Soit $G \subset \mathrm{GL}_n(\mathbb{C})$ un groupe de Lie. Alors, son algèbre de Lie est :*

$$\operatorname{Lie} G = \mathfrak{g} = \{X \in \mathrm{GL}_n(\mathbb{C}) : \forall t \in \mathbb{R},\ \exp(tX) \in G\}.$$

Ce théorème ouvre la voie à un dictionnaire de traduction entre algèbres de Lie et groupes de Lie.

Références : Pour des précisions ou des compléments, on s'en remettra à la référence [25][6], chapitres 3, 4, 5 ; voir aussi [6].

C. Exercices

C.1. Exercice (Propriétés topologiques des groupes symplectiques). *On reprend les notations de l'exercice* V-D.17.

A. Le groupe compact $\mathrm{Sp}(n)$

1. *Montrer que $\mathrm{Sp}_{2n}(\mathbb{C})$ est un sous-groupe fermé du groupe*

$$\{g \in \mathrm{GL}(2n, \mathbb{C}),\ \det(g) = \pm 1\}$$

et que $\mathrm{Sp}(n) = \mathrm{Sp}_{2n}(\mathbb{C}) \cap \mathrm{U}_{2n}(\mathbb{C})$ en est un sous-groupe compact.

[6] Les groupes de Lie à la portée de l'étudiant en licence !

2. *Montrer en utilisant l'exercice* V-D.17-*B.1 que si* X *est un vecteur propre de* $P \in \mathrm{Sp}(n)$, i.e. *si* $PX = \lambda X$, *on a aussi :* $P(J\overline{X}) = \overline{\lambda}(J\overline{X})$.

3. *Déduire en utilisant l'exercice* V-D.17-*B.3 que, pour* $P \in \mathrm{Sp}_{2n}(\mathbb{C})$ *donnée, on a :* $P \in \mathrm{Sp}(n)$ *si et seulement si* P *est* $\mathrm{Sp}(n)$*-semblable à une matrice diagonale de la forme*

$$P' = \left(\begin{array}{ccc|ccc} e^{i\theta_1} & & & & & \\ & \ddots & & & 0 & \\ & & e^{i\theta_n} & & & \\ \hline & & & e^{-i\theta_1} & & \\ & 0 & & & \ddots & \\ & & & & & e^{-i\theta_n} \end{array}\right).$$

4. *Déduire que* $\mathrm{Sp}(n)$ *est connexe par arcs et que* $\mathrm{Sp}(n) \subset \mathrm{SU}(2n)$.

5. *Montrer que le centre de* $\mathrm{Sp}(n)$ *est réduit à* $\{\pm 1\}$. *Montrer qu'il en est de même pour le centre de* $\mathrm{Sp}_{2n}(\mathbb{C})$.

B. Matrices hermitiennes symplectiques

1. *On pose* $\mathscr{H}^{\sigma}_{2n} = \mathscr{H}^{++}_{2n} \cap \mathrm{Sp}_{2n}(\mathbb{C})$. *C'est l'ensemble des matrices symplectiques hermitiennes définies positives.*

2. *Montrer que pour* $A \in \mathscr{H}^{++}_{2n}$ *on a l'équivalence entre les assertions :*

   (i) $A \in \mathrm{Sp}_{2n}(\mathbb{C})$;

   (ii) $A\,J\,\overline{A} = J$.

3. *En déduire que pour* $A \in \mathscr{H}^{\sigma}_{2n}$, *on a l'implication suivante :*

$$AX = \lambda X \quad \Longrightarrow \quad A(J\overline{X}) = \frac{1}{\lambda}(J\overline{X}).$$

4. *Montrer que pour* $A \in \mathrm{Sp}_{2n}(\mathbb{C})$ *on a l'équivalence entre les assertions :*

   (i) $A \in \mathscr{H}^{\sigma}_{2n}$;

   (ii) A *est* $\mathrm{Sp}(n)$*-semblable à une matrice réelle diagonale*

$$A' = \left(\begin{array}{ccc|ccc} \lambda_1 & & & & & \\ & \ddots & & & 0 & \\ & & \lambda_n & & & \\ \hline & & & \frac{1}{\lambda_1} & & \\ & 0 & & & \ddots & \\ & & & & & \frac{1}{\lambda_n} \end{array}\right),$$

   où $\lambda_j > 0$ *pour tout* $j \in \{1, \ldots, n\}$.

5. *En déduire les propriétés suivantes.*
   (a) *L'ensemble* $\mathscr{H}_{2n}^{\sigma}$ *est connexe par arcs.*
   (b) *Si*[7] $A \in \mathscr{H}_{2n}^{\sigma}$, *alors* $\sqrt{A} \in \mathscr{H}_{2n}^{\sigma}$ *et l'application*
   $$\sqrt{\ } : \mathscr{H}_{2n}^{\sigma} \longrightarrow \mathscr{H}_{2n}^{\sigma}$$
   est un homéomorphisme.

C. Décomposition polaire

1. *Montrer que l'on a un homéomorphisme de décomposition polaire*
   $$\delta : \mathscr{H}_{2n}^{\sigma} \times \mathrm{Sp}(n) \longrightarrow \mathrm{Sp}_{2n}(\mathbb{C}).$$
2. *En déduire que* $\mathrm{Sp}_{2n}(\mathbb{C})$ *est un sous-groupe fermé connexe de* $\mathrm{SL}(2n, \mathbb{C})$.

D. Algèbres de Lie.

1. *Montrer que l'algèbre de Lie* $\mathfrak{sp}_{2n}(\mathbb{C})$ *de* $\mathrm{Sp}_{2n}(\mathbb{C})$ *est donnée par*
   $$\mathfrak{sp}_{2n}(\mathbb{C}) = \left\{A \in \mathscr{M}_{2n}(\mathbb{C}),\ {}^tAJ + JA = 0_{2n}\right\}.$$
2. *En déduire que* $\mathfrak{sp}_{2n}(\mathbb{C}) \subset \mathfrak{sl}(2n, \mathbb{C})$ *et* $\dim \mathfrak{sp}_{2n}(\mathbb{C}) = n(2n+1)$. *Cela donne-t-il une preuve indépendante que* $\mathrm{Sp}_{2n}(\mathbb{C}) \subset \mathrm{SL}_{2n}(\mathbb{C})$ *?*

C.2. Exercice (Avatars de $\mathrm{PSL}_2(\mathbb{R})$).

L'exercice est un batifolage autour des isomorphismes exceptionnels entre $\mathrm{PSU}(1,1)$, $\mathrm{SO}_0(2,1)$ *et* $\mathrm{PSL}_2(\mathbb{R})$, *suivi d'une réflexion sur la géométrie cachée derrière l'isomorphisme* $\mathrm{PSU}(1,1) \simeq \mathrm{PSL}_2(\mathbb{R})$.

A. *On étudie dans cette partie la quadrique* $\mathscr{C}$ *associée à la forme quadratique de* $\mathbb{R}^4$ *donnée par* $q = x^2 + y^2 - z^2 - t^2$:
$$\mathscr{C} := \left\{(x, y, z, t) \in \mathbb{R}^4,\ x^2 + y^2 - z^2 - t^2 = 1\right\}.$$

1. *Montrer en utilisant le théorème de submersion que* $\mathscr{C}$ *est une sous-variété de dimension* 3 *de* $\mathbb{R}^4$.
2. *Décrire l'intersection de* $\mathscr{C}$ *avec le plan* $\{(x, y, 0, 0),\ x, y \in \mathbb{R}\}$. *En déduire que* $(1,0,0,0)$ *et* $(-1,0,0,0)$ *sont dans la même composante connexe de* $\mathscr{C}$.
3. *Montrer que* $\mathrm{O}(2,2)$ *agit transitivement sur* $\mathscr{C}$ *et que* $\mathrm{O}(2,2)/\mathrm{SO}_0(2,2)$ *agit transitivement sur les composantes connexes de* $\mathscr{C}$.
4. *À l'aide des représentants explicites de* $\mathrm{O}(2,2)/\mathrm{SO}_0(2,2)$ *dans* $\mathrm{O}(2,2)$, *déduire que* $\mathscr{C}$ *est connexe.*

[7] On rappelle que si A est une matrice hermitienne définie positive, sa racine $\sqrt{A}$ est l'unique matrice hermitienne définie positive B telle que $B^2 = A$, voir remarque VI-1.5.

B. *On note*

$$\mathrm{U}(1,1) := \{A \in \mathscr{M}_2(\mathbb{C}),\ AJA^* = J\} \text{ et } \mathrm{SU}(1,1) := U(1,1) \cap \mathrm{SL}_2(\mathbb{C}),$$

$$\text{où } J = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

1. *Montrer que* $\mathrm{U}(1,1)$ *est le sous-groupe de* $\mathrm{GL}_2(\mathbb{C})$ *des matrices de la forme*

$$\begin{pmatrix} a & \lambda\bar{b} \\ b & \lambda\bar{a} \end{pmatrix},$$

où λ *est un complexe de module 1 et* a *et* b *deux complexes tels que* $|a|^2 - |b|^2 = 1$.

2.(a) *Montrer que* $\mathrm{SU}(1,1)$ *est une sous-variété connexe de dimension* 3 *de* $\mathbb{R}^4$. *Montrer que* $\mathrm{SU}(1,1)$ *est un groupe de Lie connexe.*

(b) *Expliquer pourquoi l'algèbre de Lie* $\mathfrak{su}(1,1)$ *est incluse dans l'algèbre de Lie* $\mathfrak{u}(1,1) \cap \mathfrak{sl}_2(\mathbb{C})$.

(c) *Montrer que* $\mathfrak{u}(1,1) \cap \mathfrak{sl}_2(\mathbb{C}) = \left\{ \begin{pmatrix} iu & v+iw \\ v-iw & -iu \end{pmatrix},\ u,v,w \in \mathbb{R} \right\}$.

(d) *En déduire* $\mathfrak{su}(1,1) = \mathfrak{u}(1,1) \cap \mathfrak{sl}_2(\mathbb{C})$.

3. *En faisant agir* $\mathrm{SU}(1,1)$ *agit par conjugaison sur son algèbre de Lie, construire un isomorphisme exceptionnel* $\mathrm{PSU}(1,1) \simeq \mathrm{SO}_0(2,1)$.

On sera amené à décrire l'algèbre de Lie de $\mathrm{SO}_0(2,1)$ et à calculer sa dimension.

C. *On rappelle que* $\mathrm{PSL}_2(\mathbb{R})$ *est isomorphe à* $\mathrm{SO}_0(2,1)$. *On considère l'action par homographie de* $\mathrm{PGL}_2(\mathbb{C})$ *sur la droite projective* $\mathbb{P}^1(\mathbb{C}) = \mathbb{C} \cup \{\infty\}$ *donnée par*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az+b}{cz+d}.$$

1. *La droite des réels de* $\mathbb{C}$ *prolongée à l'infini sera appelée droite projective réelle* $\mathbb{P}^1(\mathbb{R})$ *de* $\mathbb{P}^1(\mathbb{C})$. *Montrer que le stabilisateur de* $\mathbb{P}^1(\mathbb{R})$ *dans* $\mathrm{PGL}_2(\mathbb{C})$ *est* $\mathrm{PGL}_2(\mathbb{R})$.

On pourra, à l'aide d'une méthode inouïe, utiliser la notion de repère projectif complexe et réel.

2.(a) *Définir de même la droite projective imaginaire de* $\mathbb{P}^1(\mathbb{C})$ *et trouver un élément de* $\mathrm{PGL}_2(\mathbb{C})$ *qui envoie la droite projective réelle sur la droite projective imaginaire.*

On pourra commencer par exhiber une homographie simple qui envoie $\mathbb{R}$ sur $i\mathbb{R}$.

(b) *Quel est le stabilisateur de la droite projective imaginaire ?*

3.(a) *Montrer que le stabilisateur* G *dans* $\mathrm{GL}_2(\mathbb{C})$ *de* $C := \{(z, z'), z \in i\mathbb{R}z'\} \subset \mathbb{C}^2$ *est conjugué au sous-groupe* $\mathbb{C}^* \mathrm{GL}_2(\mathbb{R})$ *engendré par les homothéties complexes et* $\mathrm{GL}_2(\mathbb{R})$.

(b) *Montrer que* C *est le cône isotrope,* i.e. *le cône des éléments isotropes, d'une forme hermitienne bien choisie et en déduire une injection* $\mathrm{SU}(1,1) \hookrightarrow G \cap \mathrm{SL}_2(\mathbb{C}) \simeq \mathrm{SL}_2(\mathbb{R})$.

(c) *Montrer que cette injection de* $\mathrm{SU}(1,1)$ *vers* $\mathrm{SL}_2(\mathbb{R})$ *est un isomorphisme.*

C.3. Exercice (Schéma de Hilbert ponctuel du plan).

On note $\mathfrak{Com}$ *la variété commutante, définie par :*

$$\mathfrak{Com} = \big\{(A, B) \in \mathscr{M}_n(\mathbb{C}),\ AB = BA\big\},$$

et $\mathfrak{Com}^0$ *l'ouvert de cette variété formé des couples* (A, B) *pour lesquels il existe un vecteur* v *dans* $\mathbb{C}^n$ *tel que* $\mathbb{C}[A, B]\cdot v = \mathbb{C}^n$ *(voir l'exercice* II-F.34*).*

1. *Soit* $f : \mathscr{M}_n(\mathbb{C}) \times \mathscr{M}_n(\mathbb{C}) \to \mathscr{M}_n(\mathbb{C})$, $(A, B) \mapsto [A, B] = AB - BA$. *Montrer que la différentielle de* f *en* (A, B) *est donnée par*

$$\mathrm{d}f_{(A,B)} : (H, K) \mapsto [A, K] + [H, B].$$

2. *On munit* $\mathscr{M}_n(\mathbb{C})$ *de la forme non dégénérée* $(M, N) = \mathrm{tr}(MN)$. *Montrer que*

$$(\mathrm{Im}\,\mathrm{d}f_{(A,B)})^{\perp} = \big\{L,\ [L, A] = [L, B] = 0\big\},$$

c'est-à-dire égale à l'intersection $\mathscr{C}_A \cap \mathscr{C}_B$ *des commutants de* A *et* B.

L'inclusion inverse est claire. Pour l'inclusion directe, on montrera que l'hypothèse « $\mathrm{tr}(LAK) = \mathrm{tr}(LKA)$, pour tout K » implique $LA = AL$, en utilisant la non dégénérescence de la forme.

3. *Soit* (A, B) *dans* $\mathfrak{Com}^0$ *et* v *comme dans la définition ci-dessus. Montrer que pour tout* w *de* $\mathbb{C}^n$, *il existe une unique matrice* L_w *de* $\mathscr{C}_A \cap \mathscr{C}_B$ *telle que* $L_w \cdot v = w$. *Puis, montrer que* $L \mapsto L\cdot v$ *et* $w \mapsto L_w$ *définissent des bijections réciproques entre* $\mathscr{C}_A \cap \mathscr{C}_B$ *et* $\mathbb{C}^n$.

4. *En déduire que la différentielle de* f *est de rang constant sur* $\mathfrak{Com}^0$.

Cet exercice est tiré de [27]. Il permet de montrer une propriété de lissité du schéma de Hilbert.

Chapitre X

Droite projective et applications

Following the line's not really my style. The Avengers, Joss Whedon, 2012.

Le héros de ce chapitre est le birapport : invariant fondamental de la géométrie projective, au même titre que le rapport de mesures algébriques pour la géométrie de la droite affine ou les angles pour la géométrie euclidienne du plan ; son acte de naissance est signé par le groupe des transformations projectives de la droite, PGL_2.

Certes, la formule que l'on apprend (apprenait ?) en terminale semble arbitraire et l'invariance par homographies y apparaît comme un miracle : classiquement, inverser le point de vue pour plonger le concept dans une plus grande généralité (le programme d'Erlangen, si l'on veut, ne lésinons pas) restitue sa majesté au birapport. Ici, on construit donc le birapport comme un invariant avant d'en déduire sa valeur.

À bon outil, monceaux d'applications. En faisant simplement varier le corps de base, on voit une diversité de situations où le birapport a des informations à fournir. Sur les réels, on effleure l'idée de perspective ; moins classique, du moins au sens des programmes universitaires, on utilise le birapport pour définir une distance sur tout convexe compact de $\mathbb{R}^n$, utile dans l'étude des billards.

Vu que la droite projective complexe est une complétion du plan affine euclidien, elle est un excellent cadre pour y faire de la géométrie des angles, où cercles et droites peuvent être traités uniformément : c'est la géométrie *conforme* ou *anallagmatique*, sous l'empire de $\mathrm{PGL}_2(\mathbb{C})$. On n'approfondit

pas le sujet ; cela sera fait au chapitre XI au prétexte de l'alternative de Steiner.

Incidemment, la droite projective complexe s'identifie aussi à une sphère. Pour mélanger réels et complexes, on trace un équateur sur la sphère et l'on se restreint à une hémisphère : c'est le plan hyperbolique, gouverné par le groupe $\mathrm{PSL}_2(\mathbb{R})$. Ce qui est remarquable ici, c'est que l'on peut comprendre le plan hyperbolique de multiples façons : avec des idées typiquement complexes, mais aussi en termes de produits scalaires dans le plan réel. On voit ainsi interagir homographies de la droite complexe et homographies du plan réel dans un élégant ballet, où la distance de Hilbert revient comme invitée d'honneur. Mais ces considérations ne seront abordées qu'au second tome.

De la sorte, deux géométries apparemment différentes s'unifient : c'est le début d'un dictionnaire, dont on verra l'étendue au chapitre XI, entre deux réalisations de la même géométrie — plutôt qu'entre deux géométries apparemment différentes —, dont l'unité semble résumée par un isomorphisme *exceptionnel* entre groupes de symétries.

On voit à l'œuvre l'idée maîtresse du programme d'Erlangen : pour décrire une géométrie, l'essentiel est de connaître son groupe de symétrie.

1. Droite projective sur un corps quelconque

1.1. Constructions de la droite projective $\mathbb{P}^1(\mathbb{K})$

Soit $\mathbb{K}$ un corps. Comme d'habitude, on a en tête les complexes, les réels et les corps finis. On note $\mathbb{P}^1(\mathbb{K})$, ou $\mathbb{P}^1$ s'il n'y a pas d'ambiguïté sur le corps $\mathbb{K}$, la droite projective sur $\mathbb{K}$. On peut la décrire de plusieurs façons :

- par définition, c'est l'ensemble des droites vectorielles de $\mathbb{K}^2$, c'est-à-dire aussi le quotient de l'ensemble des vecteurs non nuls par les homothéties, soit $(\mathbb{K}^2 \setminus \{0\})/\mathbb{K}^*$; si (x, y) est un vecteur non nul de $\mathbb{K}^2$, on notera $[x, y]$ la droite qu'il engendre ou, ce qui revient au même, l'ensemble des vecteurs (kx, ky) lorsque k parcourt $\mathbb{K}^*$;
- pour $(x, y) \in \mathbb{K}^* \times \mathbb{K}$, le quotient y/x ne dépend que de $[x, y]$, d'où une application $f_2 : \mathbb{K} \to \mathbb{P}^1(\mathbb{K})$, $t \mapsto [1, t]$; étant donné une droite vectorielle D, son antécédent éventuel par f_2 est l'intersection de D avec la droite affine d'équation $x = 1$; par suite, seule la droite $\infty = [0, 1]$ n'est pas dans l'image de f_2 ; si l'on identifie $\mathbb{K}$ et son image $U_2 = f_2(\mathbb{K})$ par f_2, on obtient une décomposition : $\mathbb{P}^1(\mathbb{K}) = \mathbb{K} \cup \{\infty\}$; c'est un exemple de la décomposition cellulaire des grassmanniennes que l'on a vue en IV-3 et en VIII-2 ; on a en fait deux applications analogues de $\mathbb{K}$ dans $\mathbb{P}^1(\mathbb{K})$:

$$\begin{array}{rcl} f_1 : \mathbb{K} & \longrightarrow & \mathbb{P}^1 \\ s & \longmapsto & [s, 1] \end{array} \quad \text{et} \quad \begin{array}{rcl} f_2 : \mathbb{K} & \longrightarrow & \mathbb{P}^1 \\ t & \longmapsto & [1, t] \end{array} ;$$

– le groupe $\mathrm{GL}_2(\mathbb{K})$ opère transitivement sur l'ensemble des droites de $\mathbb{K}^2$ (théorème de la base incomplète). Fixons comme droite de référence la droite $[1,0]$ engendrée par le premier vecteur de la base canonique, $(1,0)$; son stabilisateur est le groupe B des matrices triangulaires supérieures inversibles. L'application orbitale induit une bijection $\mathrm{GL}_2(\mathbb{K})/B \to \mathbb{P}^1(\mathbb{K})$, $gB \mapsto \mathbb{K}g(1,0)$, qui donne une nouvelle réalisation de $\mathbb{P}^1(\mathbb{K})$. Explicitons quelque peu :

$$\forall g = \begin{pmatrix} a & * \\ c & * \end{pmatrix} \in \mathrm{GL}_2(\mathbb{K}),\ \forall b = \begin{pmatrix} k & * \\ 0 & * \end{pmatrix} \in B,$$

$$gb = \begin{pmatrix} ka & * \\ kc & * \end{pmatrix} \text{ et } g\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix};$$

ainsi, $\mathbb{K}g(1,0)$ est la droite engendrée par la première colonne de la matrice g.

Si $\mathbb{K}$ est le corps des réels ou des complexes, on peut mettre une topologie sur $\mathbb{P}^1(\mathbb{K})$ de l'une des façons équivalentes suivantes (exercice II-F.23) :

– comme topologie quotient dans la réalisation $(\mathbb{K}^2 \setminus \{0\})/\mathbb{K}^*$,

– comme topologie quotient dans la réalisation $\mathrm{GL}_2(\mathbb{K})/B$,

– par une base de voisinages dans l'identification $\mathbb{P}^1 = \mathbb{K} \cap \{\infty\}$: les voisinages d'un point $t \in \mathbb{K}$ sont les ouverts de $\mathbb{K}$ contenant t et les voisinages de ∞ sont les complémentaires des ouverts ;

– par recollement : on met une topologie sur les parties U_1 et U_2 par les bijections $f_1 : \mathbb{K} \to U_1$ et $f_2 : \mathbb{K} \to U_2$; ces topologies coïncident sur l'intersection $U_1 \cap U_2$: en effet, $f_1^{-1}(U_1 \cap U_2) = \mathbb{K}^* = f_2^{-1}(U_1 \cap U_2)$, et la composée $f_2^{-1} \circ f_1 : \mathbb{K}^* \to \mathbb{K}^*$ est l'application bicontinue $s \mapsto s^{-1}$; un ouvert de $\mathbb{P}^1(\mathbb{K})$ est alors, par définition, une partie dont les intersections avec U_1 et U_2 sont ouvertes ; en fait, ces deux ouverts peuvent être utilisés comme des cartes pour mettre une structure de variété différentiable ou complexe sur $\mathbb{P}^1(\mathbb{K})$.

1.1.1. Remarque. On a vu que la droite projective, comme toute grassmannienne, est compacte (proposition II-4.3.11 et la remarque qui la suit). On dit que c'est une *compactification* de $\mathbb{K}$, car l'application f_2 est un plongement de $\mathbb{K}$ dans le compact $\mathbb{P}^1(\mathbb{K})$ comme un ouvert dense.

On peut vérifier que $\mathbb{P}^1(\mathbb{R})$ est homéomorphe à un cercle et l'on va montrer que $\mathbb{P}^1(\mathbb{C})$ est homéomorphe à une sphère de dimension 2.

1.2. Action de $\mathrm{PGL}_2(\mathbb{K})$ et birapport

1.2.1. Homographies

L'action naturelle de $\mathrm{GL}_2(\mathbb{K})$ sur $\mathbb{K}^2$ stabilise la partie formée des vecteurs non nuls et commute à l'action de $\mathbb{K}^*$. Par suite, elle induit une action sur le

quotient $(\mathbb{K}^2 \setminus \{0\})/\mathbb{K}^*$; si l'on réalise $\mathbb{P}^1(\mathbb{K})$ comme le quotient $\mathrm{GL}_2(\mathbb{K})/B$, l'action provient de la multiplication à gauche $\big(g \cdot hB = ghB$ pour g et h dans $\mathrm{GL}_2(\mathbb{K})\big)$. Une homothétie agit trivialement sur l'ensemble des droites ou, ce qui revient au même, le groupe des homothéties est contenu dans B : l'action se factorise donc en une action de $\mathrm{PGL}_2(\mathbb{K})$. Pour $g \in \mathrm{GL}_2(\mathbb{K})$, on notera $[g]$ son image dans $\mathrm{PGL}_2(\mathbb{K})$.

Que se passe-t-il en coordonnées, c'est-à-dire après transport par f_1 ou, si l'on veut, dans la réalisation $\mathbb{K} \cup \{\infty\}$? Si (x, y) est un vecteur non nul de $\mathbb{K}^2$ et

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

est un élément de $\mathrm{GL}_2(\mathbb{K})$, on a : $g(x, y) = (ax + by, cx + dy)$, si bien que pour $cx + dy \neq 0$, on a : $f_1^{-1}(ax + by, cx + dy) = (ax + by)/(cx + dy)$. Si, de plus, y n'est pas nul, alors on a : $[x, y] = f_1(x/y)$. Par suite :

$$[g] \cdot f_1\left(\frac{x}{y}\right) = f_1\left(\frac{ax + by}{cx + dy}\right).$$

On peut prolonger cette formule avec des conventions naturelles suivantes sur ∞.

1.2.2. Convention. Si a, b, c, d sont quatre éléments de $\mathbb{K}$ tels que $ad - bc \neq 0$, on pose :

$$\frac{a}{0} = \infty \text{ si } a \neq 0, \quad \frac{a\infty + b}{c\infty + d} = \frac{a}{c}.$$

En particulier, si $c = 0$, alors $a \neq 0$ et $(a\infty + b)/(c\infty + d) = \infty$; pour $a = d = 0$ et $b = c = 1$, on trouve : $1/\infty = 0$.

L'action de $\mathrm{PGL}_2(\mathbb{K})$ sur $\mathbb{P}^1(\mathbb{K})$ réalisé comme $\mathbb{K} \cup \{\infty\}$ s'écrit donc :

$$[g] \cdot t = \frac{at + b}{ct + d} \qquad \text{avec } g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{K}),\ t \in \mathbb{K} \cup \{\infty\}.$$

On peut vérifier que cette action est fidèle. Cela résulte de toute façon de la proposition 1.2.7.

1.2.3. Définition. On appelle *homographie* toute application de $\mathbb{P}^1(\mathbb{K}) = \mathbb{K} \cup \{\infty\}$ dans lui-même de la forme $t \mapsto [g] \cdot t$ pour $[g] \in \mathrm{PGL}_2(\mathbb{K})$.

1.2.4. Lemme. *Lorsque $\mathbb{K}$ est $\mathbb{R}$ ou $\mathbb{C}$, une homographie est un homéomorphisme de $\mathbb{P}^1(\mathbb{K})$ dans lui-même.*

Démonstration. Du fait que l'action de $\mathrm{GL}_2(\mathbb{K})$ sur $\mathbb{K}^2$ est continue, il résulte de la définition de la topologie de $\mathbb{P}^1(\mathbb{K})$ comme topologie quotient qu'une homographie est continue. De plus, la réciproque d'une homographie en est une, d'où la bicontinuité. □

1.2.5. Remarque. Les points fixes d'une homographie $[g]$ sont les images par la projection naturelle $\mathbb{K}^2 \setminus \{0\} \to \mathbb{P}^1(\mathbb{K})$ des vecteurs propres de g : il y en a donc entre 0 et 2.

1.2.6. Triple transitivité des homographies

1.2.7. Proposition. *L'action de* $\mathrm{PGL}_2(\mathbb{K})$ *sur* $\mathbb{P}^1(\mathbb{K})$ *est trois fois simplement transitive. Cela peut se formuler de deux façons équivalentes :*

(i) étant donné deux triplets (D_1, D_2, D_3) *et* (D'_1, D'_2, D'_3) *de droites distinctes de* $\mathbb{K}^2$*, il existe une unique homographie* $[g] \in \mathrm{PGL}_2(\mathbb{K})$ *telle que* $g(D_i) = D'_i$ *pour* $i = 1, 2, 3$ *;*

(ii) étant donné deux triplets de points distincts (z_1, z_2, z_3) *et* (z'_1, z'_2, z'_3) *dans* $\mathbb{K} \cup \{\infty\}$*, il existe une unique homographie* $h : z \mapsto (az+b)/(cz+d)$ *(où* a*,* b*,* c *et* d *sont des éléments de* $\mathbb{K}$ *tels que* $ad - bc \neq 0$*) telle que* $h(z_i) = z'_i$ *pour* $i = 1, 2, 3$.

1.2.8. Remarque. Dans le monde des groupes finis, les actions multiplement transitives sont très rares. Il résulte en effet de la classification des groupes finis simples (un monument de la pensée humaine, plusieurs milliers de pages extrêmement difficiles) que pour $n \geqslant 4$, les seuls groupes finis qui agissent n fois transitivement sur un ensemble fini sont le groupe symétrique $\mathfrak{S}_n$, le groupe alterné $\mathfrak{A}_{n+2}$ et quatre groupes finis simples appelés groupes de Mathieu (bien sûr, dans cet énoncé, il faut éviter des situations triviales comme le produit d'un groupe symétrique par un groupe quelconque qui agit trivialement).

Démonstration. On donne deux preuves, une pour chaque formulation.

(i) Notons (e_1, e_2) la base canonique de $\mathbb{K}^2$. Il suffit de démontrer la propriété avec $D_1 = \mathbb{K}e_1$, $D_2 = \mathbb{K}e_2$ et $D_3 = \mathbb{K}(e_1 + e_2)$. Soit u_i un vecteur directeur de D'_i $(i = 1, 2, 3)$; comme $D'_1 \neq D'_2$, (u_1, u_2) est une base de $\mathbb{K}^2$; on note (β_1, β_2) les coordonnées de u_3 dans cette base.

Si $[g]$ convient, on a nécessairement : $g(e_i) = \alpha_i u_i$ pour $\alpha_i \in \mathbb{K}^*$ convenables $(i = 1, 2)$. Alors, les relations $g(D_3) = D'_3$ et $g(e_1 + e_2) = \alpha_1 u_1 + \alpha_2 u_2$ imposent que (α_1, α_2) et (β_1, β_2) sont proportionnels, ce qui détermine g à une constante près, donc $[g]$ de façon unique. Inversement, l'homographie que l'on vient de décrire envoie bien D_i sur D'_i.

(ii) Il suffit de démontrer la propriété avec $z'_1 = \infty$, $z'_2 = 0$ et $z'_3 = 1$. Supposons provisoirement qu'aucun des éléments z_1, z_2 et z_3 n'est l'infini. Si h convient, z_1 est le pôle de la fraction rationnelle h et z_2 est son zéro donc $h(z)$ est de la forme $k(z-z_2)/(z-z_1)$ pour $k \in \mathbb{K}^*$ constante convenable. La dernière condition impose que $k(z_3-z_2)/(z_3-z_1) = 1$, d'où nécessairement :

$$\forall z \in \mathbb{P}^1(\mathbb{K}), \quad h(z) = \frac{z_3 - z_1}{z_3 - z_2} \times \frac{z - z_2}{z - z_1}.$$

L'unicité de h en résulte et l'on vérifie sans peine que l'homographie ainsi définie convient.

Reste à traiter le cas où l'un des z_k est ∞. Si $z_1 = \infty$, l'infini est fixe et z_2 est un zéro donc $h(z)$ est de la forme $k(z - z_2)$, puis $k = 1/(z_3 - z_2)$. Si $z_2 = \infty$, alors z_1 est un pôle et le numérateur doit être constant (pour avoir $h(\infty) = 0$) donc $h(z)$ est de la forme $k/(z - z_1)$ et $k = (z_3 - z_1)$. Si $z_3 = \infty$, z_1 est le pôle et z_2 le zéro donc $h(z) = k(z - z_1)/(z - z_2)$ et la relation $h(\infty) = k$ force $k = 1$. On constate que c'est la même formule ci-dessus, pour peu que l'on y efface les deux facteurs où apparaît le z_k infini. Elle convient dans chaque cas. □

1.2.9. Repères projectifs et homographies

On appelle *repère projectif* de $\mathbb{P}^1(\mathbb{K})$ tout triplet de points distincts. Étant donné un repère projectif (z_1, z_2, z_3), on considère l'homographie h qui envoie z_1 sur ∞, z_2 sur 0 et z_3 sur 1 : c'est une bijection de $\mathbb{P}^1(\mathbb{K})$ dans lui-même. La coordonnée d'un élément z de $\mathbb{P}^1(\mathbb{K})$ dans ce repère est, par définition, $h(z)$.

Soit (z'_1, z'_2, z'_3) un autre repère projectif de $\mathbb{P}^1(\mathbb{K})$ et soit h' l'homographie définie de façon analogue à h. Pour un point z de $\mathbb{P}^1(\mathbb{K})$, $h'(z)$ est l'image de $h(z)$ par l'homographie $h' \circ h^{-1}$: le changement de coordonnée est donc une homographie (sur les corps des réels et des complexes, c'est en particulier un homéomorphisme).

1.2.10. Orbites de quadruplets et birapport

1.2.11. Définition. Soit (z_1, z_2, z_3, z_4) quatre éléments distincts de $\mathbb{P}^1(\mathbb{K})$. Soit h l'homographie qui envoie z_1 sur ∞, z_2 sur 0 et z_3 sur 1. On définit le birapport de (z_1, z_2, z_3, z_4) par :

$$[z_1, z_2, z_3, z_4] = h(z_4).$$

1.2.12. Remarque. L'idée de cette définition relève du principe suivant : si un groupe de transformation agit n fois simplement transitivement sur une géométrie, il est intéressant de regarder les orbites de $n+1$ points. À tout le moins, on obtient un paramétrage des orbites de $n+1$ points par (une partie de) l'espace.

Voici deux situations analogues : les vecteurs et les angles. Le groupe des translations (resp. des rotations) agit simplement transitivement sur les points d'un espace affine (resp. le cercle unité) ; on étudie les orbites de couples de points, ce qui donne lieu aux vecteurs (resp. aux angles). Ici, comme le groupe qui agit est commutatif, la relation de Chasles permet de définir une structure de groupe sur les orbites de couples : pour les vecteurs, on retrouve une partie de la structure affine que l'on connaît déjà mais, pour les angles, c'est une nouvelle opération particulièrement utile.

L'intérêt de cette définition abstraite est de fournir une preuve sans calcul de la propriété principale du birapport, l'invariance par homographies. De plus, la formule (iii) ci-dessous n'apparaît plus comme une miraculeuse atrocité mais comme une nécessité harmonieuse.

1.2.13. Proposition.

(i) *Le birapport est invariant par homographies.*

(ii) *Deux quadruplets* (z_1, z_2, z_3, z_4) *et* (z'_1, z'_2, z'_3, z'_4) *de points distincts ont le même birapport si et seulement si il existe une homographie* $g \in \mathrm{PGL}_2(\mathbb{K})$ *telle que* $g(z_i) = z'_i$ $(1 \leqslant i \leqslant 4)$.

(iii) *On a, pour* z_1, z_2, z_3, z_4 *distincts dans* $\mathbb{P}^1(\mathbb{K})$ *:*

$$[z_1, z_2, z_3, z_4] = \frac{z_3 - z_1}{z_3 - z_2} \times \frac{z_4 - z_2}{z_4 - z_1}.$$

1.2.14. Remarque. Constatons que si l'on fixe trois des variables, le birapport est une homographie en la quatrième. Si une des variables vaut ∞, la valeur du birapport dans la formule (iii) se calcule avec les conventions sur le calcul avec ∞ du paragraphe 1.2.1.

Le membre de droite de (iii) admet un sens lorsque les z_i sont des indéterminées, qui permet de *définir le birapport* $[z_1, z_2, z_3, z_4]$ dans ce contexte.

Enfin, si l'on spécialise les indéterminées à un élément de $\mathbb{P}^1(\mathbb{K})$, le birapport de la spécialisation est bien la spécialisation du birapport.

Démonstration. (i) Soit g une homographie, $(z_i)_{1 \leqslant i \leqslant 4}$ un quadruplet de points distincts de $\mathbb{P}^1(\mathbb{K})$ et $z'_i = g(z_i)$ pour tout i. On doit montrer que $[z'_1, z'_2, z'_3, z'_4] = [z_1, z_2, z_3, z_4]$. Soit donc h (resp. h') l'unique homographie qui envoie (z_1, z_1, z_3) (resp. (z'_1, z'_2, z'_3)) sur $(\infty, 0, 1)$. Par construction,

on a : $h(z_i) = h' \circ g(z_i)$ pour $i = 1, 2, 3$; comme l'action des homographies est *simplement* transitive sur les triplets, on en déduit que $h = h' \circ g$. Or, par définition du birapport, on a alors :

$$[z_1, z_2, z_3, z_4] = h(z_4) = h' \circ g(z_4) = h'(z'_4) = [z'_1, z'_2, z'_3, z'_4].$$

(ii) Soit h l'homographie telle que $h(z_1) = \infty$, $h(z_2) = 0$ et $h(z_3) = 1$. On définit de même h'. Alors, le birapport des deux quadruplets est : $h(z_4) = h'(z'_4)$. Par suite, $g = h'^{-1} \circ h$ envoie le premier quadruplet sur le second.

(iii) Il suffit d'appliquer à z_4 la formule obtenue dans la deuxième preuve de la proposition 1.2.7. □

1.2.15. Remarque. On trouvera des variantes et une interprétation dans le paragraphe 2.1.

1.2.16. Proposition (Birapport comme invariant total). *Soit* $(\mathbb{P}^1)^4_{\mathrm{reg}}$ *l'ensemble des quadruplets de points deux à deux distincts de* $\mathbb{P}^1$. *Le groupe* $\mathrm{PGL}_2(\mathbb{K})$ *agit diagonalement sur* $(\mathbb{P}^1)^4_{\mathrm{reg}}$.

(i) *Le birapport induit une bijection*

$$\overline{\lambda} : (\mathbb{P}^1)^4_{\mathrm{reg}} / \mathrm{PGL}_2(\mathbb{K}) \longrightarrow \mathbb{P}^1(\mathbb{K}) \setminus \{0, \infty, 1\},$$

$$\mathrm{PGL}_2(\mathbb{K}) \cdot (z_1, z_2, z_3, z_4) \longmapsto [z_1, z_2, z_3, z_4].$$

(ii) *Lorsque* $\mathbb{K}$ *est le corps des réels ou des complexes, on munit le quotient* $(\mathbb{P}^1)^4_{\mathrm{reg}} / \mathrm{PGL}_2(\mathbb{K})$ *de la topologie quotient (définition* II-D.1.2*). Alors*

$$\overline{\lambda} : (\mathbb{P}^1)^4_{\mathrm{reg}} / \mathrm{PGL}_2(\mathbb{K}) \to \mathbb{P}^1(\mathbb{K}) \setminus \{0, \infty, 1\}$$

est un homéomorphisme.

Démonstration. (i) C'est une reformulation de la proposition 1.2.13. Tout d'abord, le birapport d'un quadruplet (z_1, z_2, z_3, z_4) formé de points distincts ne peut être égal à ∞, 0 ou 1, car, avec les notations de la proposition 1.2.13, c'est $h(z_4)$ et h est injective. Pour $z \in \mathbb{P}^1(\mathbb{K}) \setminus \{\infty, 0, 1\}$, on a : $z = \lambda(\infty, 0, 1, z)$, donc z est l'image par $\overline{\lambda}$ de l'orbite de $(\infty, 0, 1, z)$ et $\overline{\lambda}$ est surjective. Par la proposition 1.2.13 (ii), deux quadruplets ont le même birapport si et seulement si ils sont dans la même orbite, ce qui donne l'injectivité de $\overline{\lambda}$.

(ii) L'application $\overline{\lambda} : (\mathbb{P}^1)^4_{\mathrm{reg}} / \mathrm{PGL}_2(\mathbb{K}) \to \mathbb{P}^1 \setminus \{0, 1, \infty\}$ est bien définie, bijective d'après (i) et continue, car elle provient de l'application continue $\lambda : (\mathbb{P}^1)^4_{\mathrm{reg}} \to \mathbb{P}^1$, $(z_1, z_2, z_3, z_4) \to [z_1, z_2, z_3, z_4]$ par passage au quotient.

D'autre part, la réciproque de cette application est la composée des applications continues suivantes :

$$\begin{aligned} s : \mathbb{P}^1 \setminus \{\infty, 0, 1\} &\longrightarrow (\mathbb{P}^1)^4_{\mathrm{reg}} \\ z &\longmapsto (\infty, 0, 1, z) \end{aligned}$$

et

$$\begin{aligned} \pi : (\mathbb{P}^1)^4_{\mathrm{reg}} &\longrightarrow (\mathbb{P}^1)^4_{\mathrm{reg}} / \mathrm{PGL}_2(\mathbb{K}) \\ (z_1, z_2, z_3, z_4) &\longmapsto \mathrm{PGL}_2(\mathbb{K}) \cdot (z_1, z_2, z_3, z_4). \end{aligned}$$

□

1.2.17. Remarque. Pour $\mathbb{K}$ le corps des réels ou des complexes, on voit ainsi $\mathbb{P}^1(\mathbb{K})$ comme une *compactification* du quotient $(\mathbb{P}^1)^4_{\mathrm{reg}} / \mathrm{PGL}_2(\mathbb{K})$. On y reviendra au second tome.

1.2.18. Permutation des variables du birapport

Le groupe symétrique $\mathfrak{S}_4$ agit sur les lettres z_1, z_2, z_3, z_4 ; observons son effet sur le birapport.

Tout d'abord, le groupe de Klein agit trivialement sur le birapport. Le groupe de Klein est formé des doubles transpositions et de l'identité, il est isomorphe à $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, on vérifie immédiatement que l'on a :

$$[z_1, z_2, z_3, z_4] = [z_2, z_1, z_4, z_3] = [z_3, z_4, z_1, z_2] = [z_4, z_3, z_2, z_1].$$

Il en résulte que par permutation des variables, le birapport prend au plus $24/4 = 6$ valeurs possibles — en effet, le groupe de Klein est contenu dans le stabilisateur du birapport, quand on le considère comme un élément du corps $\mathscr{F}$ des fractions rationnelles en quatre indéterminées.

1.2.19. Remarque. Le groupe de Klein est un sous-groupe distingué dans $\mathfrak{S}_4$ (c'est la réunion de deux classes de conjugaison), donc le quotient est un groupe d'ordre 6 ; on peut bien croire que ce quotient n'est pas abélien et en déduire qu'il est isomorphe à $\mathfrak{S}_3$.

Mais, le birapport permet de le démontrer sans calcul : en effet, le stabilisateur de z_4 dans $\mathscr{F}$ est naturellement $\mathfrak{S}_3$, par permutation de z_1, z_2 et z_3 (du moins, il le contient) ; de plus, l'intersection de $\mathfrak{S}_3$ avec le groupe de Klein est triviale ; par conséquent, $\mathfrak{S}_3$ s'injecte dans le quotient de $\mathfrak{S}_4$ par le groupe de Klein. Par égalité des cardinaux, cette injection est un isomorphisme et l'on peut même affirmer que $\mathfrak{S}_4$ est le produit semi-direct de $\mathfrak{S}_3$ par le groupe de Klein grâce à la proposition II-5.2.1.

D'après la remarque ci-dessus, pour étudier l'action de $\mathfrak{S}_4$ sur le birapport, il suffit d'étudier l'action des permutations des trois premières variables ; de

plus, comme les triplets de points distincts sont tous dans la même orbite sous $\mathrm{PGL}_2(\mathbb{K})$, il suffit d'étudier le cas où $(z_1, z_2, z_3) = (\infty, 0, 1)$. Les calculs sont très faciles et seront omis.

1.2.20. Lemme. *L'action de $\mathfrak{S}_3$ par permutation de $\infty, 0, 1$ sur le birapport $[\infty, 0, 1, z_4]$ est :*

$$\begin{array}{llll} \mathrm{Id} & [\infty,0,1,z_4] = z_4 & (\infty 0) & [0,\infty,1,z_4] = \dfrac{1}{z_4} \\ (\infty 01) & [0,1,\infty,z_4] = \dfrac{z_4-1}{z_4} & (01) & [\infty,1,0,z_4] = -(z_4-1) \\ (\infty 10) & [1,\infty,0,z_4] = -\dfrac{1}{z_4-1} & (\infty 1) & [1,0,\infty,z_4] = \dfrac{z_4}{z_4-1}. \end{array}$$

Bien que l'on n'en voie guère l'origine, on constate la relation :

$$[\infty, 0, 1, z_4] + [\infty, 1, 0, z_4] = 1,$$

et le produit sur la première colonne donne :

$$[\infty, 0, 1, z_4] \times [0, 1, \infty, z_4] \times [1, \infty, 0, z_4] = z_4 \times \frac{z_4 - 1}{z_4} \times \frac{-1}{z_4 - 1} = -1.$$

Grâce à la triple transitivité des homographies et à l'invariance du birapport, on en déduit la propriété suivante.

1.2.21. Lemme. *Soit z_1, z_2, z_3, z_4 quatre éléments distincts de $\mathbb{P}^1(\mathbb{K})$. Alors :*

$$[z_1, z_2, z_3, z_4] + [z_1, z_3, z_2, z_4] = 1$$

$$\textit{et} \quad [z_1, z_2, z_3, z_4] \times [z_2, z_3, z_1, z_4] \times [z_3, z_1, z_2, z_4] = -1.$$

La première relation est utilisée pour démontrer le théorème de Ptolémée, voir l'exercice C.8.

1.2.22. Division harmonique et équianharmonique

On vient de voir que le birapport prend, génériquement, six valeurs lorsque l'on permute les quatre points. Il peut arriver que, parmi ces valeurs, plusieurs coïncident. Par transitivité, on peut se limiter aux quadruplets $\{\infty, 0, 1, z\}$. Les six valeurs sont des homographies qui forment un sous-groupe de $\mathrm{PGL}_2(\mathbb{K})$ isomorphes à $\mathfrak{S}_3$: l'ensemble des valeurs qui coïncident avec z est nécessairement un sous-groupe. Les possibilités sont les suivantes.

- *Le stabilisateur est d'ordre* 2. Il y a trois possibilités : soit $z = 1/z$, auquel cas $z = -1$; soit $z = 1 - z$, auquel cas $z = 1/2$; soit $z = z/(z-1)$, auquel cas $z = 2$.
  Les trois valeurs -1, $1/2$ ou 2 sont dans la même orbite sous $\mathfrak{S}_3$:
  $$[\infty, 0, 1, -1] = -1, \quad [\infty, 1, 0, -1] = 2, \quad [1, 0, \infty, -1] = \frac{1}{2}\cdot$$
  Un quadruplet ordonné (z_1, z_2, z_3, z_4) dont le birapport vaut -1 est appelé une *division harmonique*. En oubliant l'ordre des points, on peut décrire un tel quadruplet de la façon suivante : si l'un des points est à l'infini, un des trois qui restent est au milieu des deux autres.
- *Le stabilisateur est d'ordre* 3. Il y a une seule possibilité : $z = (z-1)/z$, ce qui équivaut à $z = -1/(z-1)$, ou encore à $z^2 - z + 1 = 0$, c'est-à-dire que $-z$ est une racine primitive cubique de 1.
  Supposons que $\mathbb{K}$ contienne des racines cubiques j et j^2 de l'unité distinctes de 1 ; cela revient à dire que -3 est un carré dans $\mathbb{K}$. Alors, les scalaires $-j$ et $-j^2$ sont dans la même orbite sous $\mathfrak{S}_3$, puisque l'on a :
  $$[\infty, 0, 1, -j] = -j, \quad [0, \infty, 1, -j] = -j^2.$$
  Un quadruplet dont le birapport vaut $-j$ ou $-j^2$ est appelé par Pierre Samuel dans [31] une *division équianharmonique*. Si l'un des points est à l'infini, les trois autres forment un triangle équilatéral (du moins, dans $\mathbb{C}$). Quels théorèmes de géométrie sont capturés par cette configuration ? On en verra une application de nature plutôt combinatoire au second volume.

Pour référence ultérieure, on pose la définition suivante.

1.2.23. Définition. Soit $\mathbb{K}$ un corps. On appelle *partie harmonique* (resp. *partie équianharmonique*) une partie à quatre éléments $\{z_1, z_2, z_3, z_4\}$ de la droite $\mathbb{P}^1(\mathbb{K})$ dont le birapport, dans n'importe quel ordre, appartient à $\{-1, 2, 1/2\}$ (resp. à $\{-j, -j^2\}$, où j est une racine primitive cubique de l'unité dans $\mathbb{K}$).

1.3. Formule des six birapports (d'après Perrin, Audin et Lass)

Il existe une autre formule évidente à vérifier mais dont l'origine est moins claire, que nous appellerons comme le fait Michèle Audin la *formule des six birapports de Perrin*, voir [4] et l'annexe du rapport sur l'enseignement de la géométrie dans [18]. Rappelons que le birapport de quatre indéterminées est défini par la formule (iii) de la proposition 1.2.13.

1.3.1. Proposition. *Soit huit indéterminées a, b, c, d, p, q, r, s. On a :*
$$[a, b, c, d] \times [a, r, d, p] \times [a, q, p, c] \times [s, c, q, b] \times [s, p, r, q] \times [s, d, b, r] = 1.$$

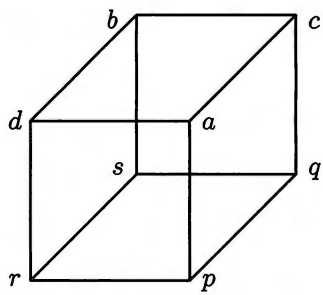
1.3.2. Remarque. Cette formule s'utilise généralement dans $\mathbb{P}^1(\mathbb{C})$ de la façon suivante, fondée sur le critère de cocyclicité/alignement en termes de birapport : on s'arrange pour être sûr que cinq des six birapports sont réels ; on en déduit que le sixième est réel.

Pour cet usage, la seule chose qui compte est que le produit des six birapports soit une constante réelle non nulle. Pour éviter des calculs ennuyeux mais triviaux, la démonstration donnera une version faible de la proposition.

1.3.3. Un cube pour six birapports

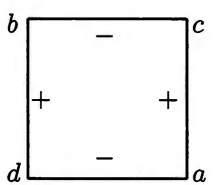
Malgré les apparences, cette formule est très simple. Voici un moyen raisonnable pour la retrouver en regardant un cube. Pourquoi un cube ? C'est une façon simple de grouper de six façons différentes par quatre lettres choisies parmi huit. En effet, un cube a huit sommets et six faces de quatre sommets[1].

Bref. On numérote donc les sommets d'un cube par nos indéterminées :



La formule des six birapport se paraphrase ainsi : on écrit un birapport par face, le produit sur les faces vaut 1.

Notre collègue Bodo Lass nous a montré la figure suivante. Si l'on ne prête pas (encore, ou trop) attention aux signes, on peut coder le birapport $[a, b, c, d]$ par un carré affecté de signes :



$$\frac{c-a}{d-a} \times \frac{d-b}{c-b}$$

[1] Il n'aura pas échappé aux fanatiques de polyèdres réguliers que le groupe des isométries du cube est justement le groupe $\mathfrak{S}_4$ dont on parlait naguère... Mais quel lien avec la formule des six birapports, je vous demande un peu ?

Un signe + (resp. −) sur une arête indique que la différence des variables attachées aux sommets de l'arête apparaît au numérateur (resp. au dénominateur) du birapport.

La formule devient évidente au signe près si l'on regarde le cube en perspective qui suit, dans lequel la face « du dessous » *pqsr* a été éclatée et se trouve représentée comme l'extérieur du carré.

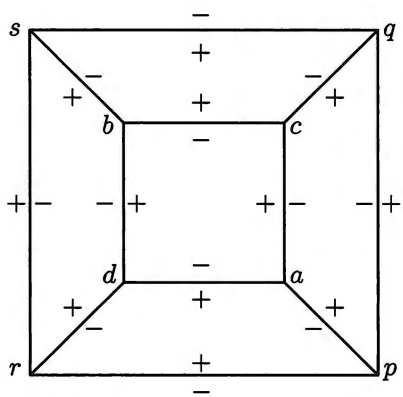


Figure 1.1. *Un cube pour la formule des six birapports*

Dans le produit des six birapports associés aux faces, chaque différence de deux indéterminées, correspondant à chaque arête, intervient deux fois, une fois au numérateur et une fois au dénominateur. On peut donc affirmer sans calculs que le produit vaut -1 ou 1.

2. Birapport dans un plan projectif réel

Ce paragraphe demande une certaine familiarité avec la géométrie projective. Les éléments les plus rudimentaires sont résumés dans l'annexe B. Les livres de Michèle Audin [4] ou de Pierre Samuel [31] sont des références plus recommandables.

2.1. Le birapport comme rapport de rapports

On a défini le birapport de quatre éléments de $\mathbb{P}^1(\mathbb{K})$. En voici quelques variantes et une interprétation comme « rapport de rapports ».

2.1.1. Birapport de quatre points sur une droite projective

Étant donné quatre points sur une droite projective, on souhaite définir leur birapport. Pour cela, on fixe un repère projectif de la droite et l'on calcule

le birapport des quatre coordonnées. Le résultat obtenu ne dépend pas du repère, car un changement de repère est une homographie et le birapport est invariant par homographie.

2.1.2. Birapport de quatre droites sur un pinceau

2.1.3. Définition. Un *pinceau de droites* est l'ensemble des droites qui passent par un point fixé. Si p est le point en question, on note p^* le pinceau correspondant.

2.1.4. Proposition. *On se donne d_1, d_2, d_3 et d_4 quatre droites d'un pinceau. Soit δ et δ' deux droites qui les coupent toutes les quatre et, pour i compris entre 1 et 4, soit m_i (resp. m'_i) l'intersection de d_i et δ (resp. δ'). Alors, les birapports $[m_1, m_2, m_3, m_4]$ et $[m'_1, m'_2, m'_3, m'_4]$ sont égaux.*

Cette proposition donne un sens à la définition suivante.

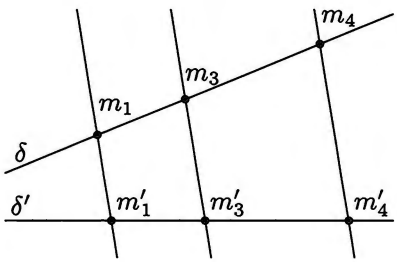
2.1.5. Définition. On appelle birapport de quatre droites d_1, d_2, d_3, d_4 appartenant à un même pinceau et l'on note $[d_1, d_2, d_3, d_4]$ le birapport $[m_1, m_2, m_3, m_4]$, où le point m_i est l'intersection de d_i avec une droite quelconque δ $(i = 1, \ldots, 4)$.

Démonstration. Quitte à changer de repère, on peut supposer que le point commun à toutes les droites est l'origine. Fixons deux bases de $\mathbb{R}^2$, correspondant à des systèmes de coordonnées (x, y) et (x', y'), telles que δ (resp. δ') soit la droite d'équation $y = 1$ (resp. $y' = 1$). On peut calculer le birapport des m_i comme le birapport des coordonnées x_i/y_i, de même pour les m'_i. Or, x'_i et y'_i sont, par un simple changement de base dans $\mathbb{R}^2$, des combinaisons linéaires de x_i et y_i ; par suite, x'_i/y'_i est une fonction homographique de x_i/y_i. On peut conclure grâce à l'invariance du birapport par homographie dans $\mathbb{P}^1(\mathbb{K})$. □

2.1.6. Remarque. Dans un plan projectif, un pinceau est naturellement une droite projective. Si V est un espace vectoriel de dimension 3, une droite du plan projectif $\mathbb{P}(V)$ est le noyau d'une forme linéaire ℓ ; elle s'identifie au point $\mathbb{K}\ell$ de $\mathbb{P}(V^*)$. Si le point p est (l'image dans $\mathbb{P}(V)$ de) la droite $\mathbb{K}v$ de V, alors $\operatorname{Ker}\ell$ appartient à p^* si et seulement si $\ell(v) = 0$. Cette contrainte sur la forme linéaire ℓ définit un plan vectoriel dans V^*, c'est-à-dire une droite projective dans $\mathbb{P}(V^*)$. On peut vérifier que le birapport de la proposition est bien le birapport défini ci-dessus sur la droite projective p^*.

2.1.7. Préservation du rapport de rapports par projection centrale

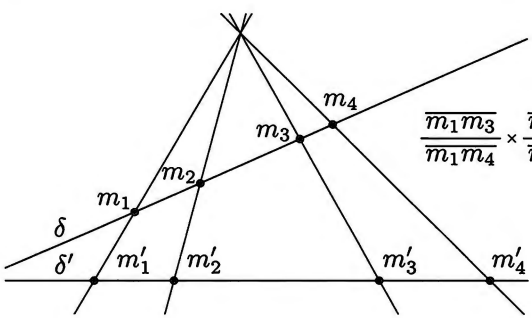
Le théorème de Thalès exprime que le rapport de mesures algébriques est invariant par projection parallèle (projection sur une droite, parallèlement à une autre droite).



$$\frac{\overline{m_1m_3}}{\overline{m_1m_4}} = \frac{\overline{m'_1m'_3}}{\overline{m'_1m'_4}}.$$

Figure 2.1. *Invariance du rapport par projection parallèle : Thalès*

Remplaçons la projection parallèle par une projection centrale, de centre p sur la droite δ' : c'est l'application qui envoie un point m du plan distinct de p sur l'intersection de δ' et de la droite (mp) (ou le point à l'infini de δ' en cas de parallélisme). On s'intéresse à la restriction de la projection à une droite δ.



$$\frac{\overline{m_1m_3}}{\overline{m_1m_4}} \times \frac{\overline{m_2m_4}}{\overline{m_2m_3}} = \frac{\overline{m'_1m'_3}}{\overline{m'_1m'_4}} \times \frac{\overline{m'_2m'_4}}{\overline{m'_2m'_3}}.$$

Figure 2.2. *Invariance du birapport par projection centrale*

Les rapports n'ont plus de raison d'être égaux. Mais, ajoutons un quatrième point : la proposition 2.1.4 exprime que les birapports $[m_1, m_2, m_3, m_4]$ et $[m'_1, m'_2, m'_3, m'_4]$ sont égaux.

2.1.8. Remarque. Plaçons-nous dans un plan projectif. Quitte à appliquer une homographie (du plan projectif) convenable, on peut supposer que la droite $(m_2m'_2)$ est la droite à l'infini. Mais alors, le birapport $[m_1, m_2, m_3, m_4]$ est le rapport $\overline{m_1m_3}/\overline{m_1m_4}$, *idem* pour $[m'_1, m'_2, m'_3, m'_4]$.

De plus, dans le plan affine correspondant, les droites (m_1m_1'), (m_3m_3') et (m_4m_4') sont parallèles : d'après le théorème de Thalès, les rapports $\overline{m_1m_3}/\overline{m_1m_4}$ et $\overline{m_1'm_3'}/\overline{m_1'm_4'}$ sont égaux, ce qui démontre la préservation du birapport par projection centrale. On en déduit une autre preuve de la proposition 2.1.4.

2.1.9. Birapport de quatre points sur une conique

Soit m_i $(1 \leqslant i \leqslant 4)$ quatre points sur une conique non dégénérée du plan. Soit m un point de la conique distinct des m_i. Notons d_i la droite (mm_i) $(1 \leqslant i \leqslant 4)$: alors le birapport $[d_1, d_2, d_3, d_4]$ ne dépend que des m_i et pas de m (voir [4], corollaire VII.4.3). On l'appelle birapport des quatre points, on le note $[m_1, m_2, m_3, m_4]$.

2.2. Aparté : PGL_2 et PSL_2

Dans la suite, nous allons observer la droite projective sur des corps différents. Selon le corps $\mathbb{K}$ de base, il est parfois utile de faire intervenir le groupe $\mathrm{PSL}_2(\mathbb{K})$, quotient du groupe $\mathrm{SL}_2(\mathbb{K})$ par les homothéties qu'il contient : soit parce qu'il s'identifie à $\mathrm{PGL}_2(\mathbb{K})$ — c'est le cas pour les complexes — , soit au contraire parce qu'il s'y distingue — c'est le cas pour les réels, où il est le stabilisateur de l'équateur dans le sphère de Riemann et de la demi-sphère, connue comme une réalisation du plan hyperbolique ; voir le second tome. Avant ce défilé magnifique[2], précisons les relations entre ces deux groupes.

Lemme. *Soit $\mathbb{K}$ un corps. Le déterminant induit un isomorphisme*

$$\mathrm{PGL}_2(\mathbb{K})/\mathrm{PSL}_2(\mathbb{K}) \xrightarrow{\simeq} \mathbb{K}^*/\mathbb{K}^{*2}.$$

Ainsi, si tout scalaire a une racine carrée, $\mathrm{PSL}_2(\mathbb{K})$ est égal à $\mathrm{PGL}_2(\mathbb{K})$; sinon, $\mathrm{PSL}_2(\mathbb{K})$ est un sous-groupe distingué d'indice 2 de $\mathrm{PGL}_2(\mathbb{K})$.

Démonstration. Comme le déterminant d'une homothétie est un carré, si g est un élément de $\mathrm{GL}_2(\mathbb{K})$ et $[g]$ est l'homographie associée, la classe de $\det(g)$ modulo les carrés de $\mathbb{K}^*$ ne dépend que de $[g]$. En d'autres termes, la composée des morphismes $\varphi : \mathrm{GL}_2(\mathbb{K}) \xrightarrow{\det} \mathbb{K}^* \to \mathbb{K}^*/\mathbb{K}^{*2}$ se factorise à travers $\mathrm{PGL}_2(\mathbb{K})$. Chacun de ces morphismes est surjectif, donc leur composée φ l'est. D'autre part, $[g]$ appartient au noyau de φ si et seulement si $\det(g)$ est un carré, disons λ^2. Dans ce cas, $\lambda^{-1}g$ appartient à $\mathrm{SL}_2(\mathbb{K})$ et $[g] = [\lambda^{-1}g]$; autrement dit, $[g] \in \mathrm{PSL}_2(\mathbb{K})$. Par suite, le noyau de φ est $\mathrm{PSL}_2(\mathbb{K})$, ce qui permet de conclure. □

[2] Que l'on ne nous intente pas pour autant le procès de faire le culte du corps !

2.3. Distance de Hilbert sur un ouvert convexe

La présentation ci-dessous est fortement inspirée de [21].

Soit K un ouvert convexe du plan qui ne contient aucune droite. On va définir une distance d sur K. Soit x et y deux points distincts de K : l'intersection de la droite (xy) et de K est convexe, c'est un intervalle ouvert $]x_0y_0[$ où l'un des deux points x_0 et y_0 au plus peut être à l'infini car K ne contient pas de droite. On nomme ses extrémités de sorte à trouver dans l'ordre : x_0, x, y, y_0.

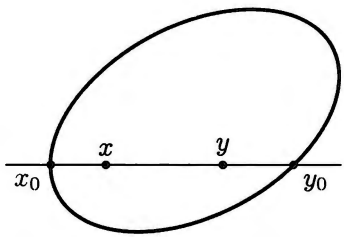


Figure 2.3. *Définition de la distance de Hilbert*

On définit[3] :

$$d(x,y) = \ln[x,y,y_0,x_0] = \ln \frac{x_0y}{x_0x} \frac{y_0x}{y_0y}.$$

(Dans le membre de droite, ce sont des distances. Si l'un des points x_0 et y_0 est à l'infini, on oublie le rapport de distances qui le contient, conformément aux conventions sur l'infini.) Le nombre obtenu est un réel strictement positif puisque $x_0y > x_0x$ et $y_0x > y_0y$. De plus, par symétrie des variables, on a : $d(x,y) = d(y,x)$.

Lorsque $x = y$, on choisit une droite contenant x, et l'on appelle x_0 et y_0 ses intersections avec le bord de K ; la formule ci-dessus donne naturellement : $d(x,x) = 0$.

2.3.1. Proposition.

L'application d ainsi définie est une distance.

Démonstration. D'après les remarques précédant la proposition, il n'y a plus qu'à prouver l'inégalité triangulaire. Pour simplifier l'exposition, on suppose que K est borné. Soit x, y et z trois points distincts de l'intérieur de K. On suppose provisoirement que z n'appartient pas à la droite (xy). On définit x_0 et y_0 comme ci-dessus et l'on note x_1 et z_1 (resp. y_2 et z_2)

[3] L'ordre des lettres, dû à la convention classique de la proposition 1.2.13, n'est pas très agréable. Ici, on aurait préféré obtenir : $[x_0,x,y,y_0] = \ln\big((x_0y/x_0x)\times(y_0x/y_0y)\big)$. Tant pis.

les intersections de la droite (xz) (resp. (yz)) avec le bord de K, de sorte à trouver dans l'ordre x_1, x, z et z_1 (resp. y_2, y, z et z_2). Soit w l'intersection[4] des droites (x_1z_2) et (y_2z_1). Soit z' l'image de z par la projection centrale[5] de centre w sur (xy), intersection de (wz) et (xy). Soit enfin x' et y' les intersections respectives de (x_1z_2) avec (xy) et de (y_2z_1) avec (xy). (Voir figure 2.4.)

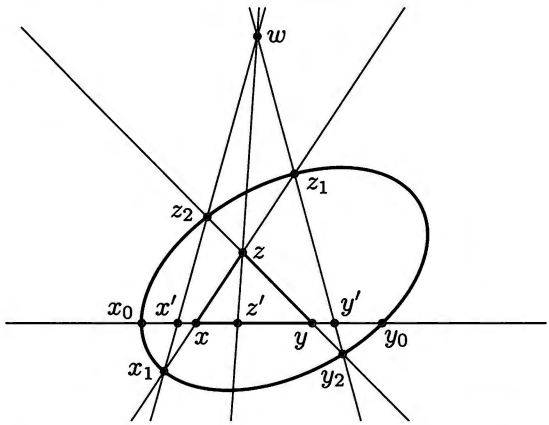


Figure 2.4. *Inégalité triangulaire pour la distance de Hilbert*

On observe sur la figure que x' est entre x_0 et x et que y' est entre y et y_0 : démontrons l'assertion concernant x' par acquit de conscience.

> Vu l'ordre des points sur les droites (yz) et (xz), on déduit que z_2 appartient au même demi-plan de frontière (xy) que z, au contraire de x_1. Par suite, le segment $[x_1z_2]$ coupe (xy). Par convexité, l'intersection x' appartient à l'adhérence de K. Or, l'intersection de (xy) et de cette adhérence est $[x_0y_0]$: ainsi, x' est entre x_0 et y_0.
>
> Vu l'ordre des points sur la droite (yz), on peut dire que z_2 n'est pas dans le même demi-plan de frontière (xz) que y ; quant à x_1, il est sur la droite (xz). Par suite, x', qui est strictement entre x_1 et z_2, n'est pas non plus dans le même demi-plan de frontière (xz) que y. Par suite, x' n'est pas entre x et y_0, ou encore : $x' \in [x_0x]$. Enfin, si $x' = x$, alors $z_1 = z_2 = z$ ou quelque absurdité comparable.

Par invariance du birapport par la projection centrale[6] de centre w, on a :

$$d(x,z) = \ln[x,z,z_1,x_1] = \ln[x,z',y',x']$$
$$\text{et} \quad d(z,y) = \ln[z,y,y_2,z_2] = \ln[z',y,y',x'],$$

[4]Si elle existe.

[5]Si les droites (x_1z_2) et (y_2z_1) sont parallèles, il s'agit de la projection sur (xy) parallèlement à leur direction.

[6]Même remarque qu'à la note précédente.

donc la somme s'écrit et se simplifie ainsi :

$$\begin{aligned} d(x,z)+d(z,y) &= \ln[x,z',y',x'] + \ln[z',y,y',x'] \\ &= \ln \frac{x'z'}{x'x}\,\frac{y'x}{y'z'}\,\frac{x'y}{x'z'}\,\frac{y'z'}{y'y} = \ln[x,y,y',x']. \end{aligned}$$

Fixons un repère affine sur la droite (xy) et identifions un point et son abscisse. Les inégalités $x_0 \leqslant x'$ et $y' \leqslant y_0$ donnent, puisque la fonction $t \mapsto (y-t)/(x-t)$ est strictement croissante sur chaque intervalle où elle est définie :

$$\begin{aligned} d(x,z)+d(z,y) &= \ln[x,y,y',x'] = \ln \frac{y-x'}{x-x'}\,\frac{x-y'}{y-y'} \\ &\geqslant \ln \frac{y-x_0}{x-x_0}\,\frac{x-y_0}{y-y_0} = d(x,y). \end{aligned}$$

□

L'exercice C.2 montre que si K est strictement convexe et x, y, z sont trois points de K, on a : $d(x,z)+d(z,y) = d(x,y)$ si et seulement si z appartient au segment $[xy]$; en revanche, si K n'est pas strictement convexe, il peut y avoir égalité alors que z n'est pas sur le segment.

3. La sphère de Riemann

Par définition, la *sphère de Riemann* est la droite complexe $\mathbb{P}^1(\mathbb{C})$. La construction du paragraphe 3.1 justifie cette dénomination et donne un contenu géométrique aux applications f_1 et f_2 du paragraphe 1.1.

3.1. Projection stéréographique

3.1.1. Définition de la projection stéréographique

Partons de la réalisation de $\mathbb{P}^1(\mathbb{C})$ comme $\mathbb{C} \cup \{\infty\}$. Considérons l'espace euclidien standard $\mathbb{R}^3$, dans lequel les coordonnées définies par la base canonique seront notées (x,y,z). Réalisons, dans $\mathbb{R}^3$, le plan complexe $\mathbb{C}$ comme le plan équatorial, *i.e.* le plan d'équation $z=0$. Soit $\mathbf{S}^2$ la sphère centrée à l'origine et de rayon 1. On note N le point $(0,0,1)$ — le pôle nord — et S le point $(0,0,-1)$ — le pôle sud. On appelle projection stéréographique de pôle N l'application $p : \mathbf{S}^2 \setminus \{N\} \to \mathbb{C}$ définie ainsi : étant donné un point $M=(x,y,z)$ de la sphère, on considère l'intersection $P=(u,v,0)$ de la droite (MN) et du plan équatorial ; on pose alors : $p(M) = u+iv$. Cela a un sens, car la droite (MN) n'est pas parallèle au plan (en effet, $z<1$). On prolonge p à $\mathbf{S}^2$ en posant : $p(N)=\infty$, ce qui constitue un prolongement par continuité.

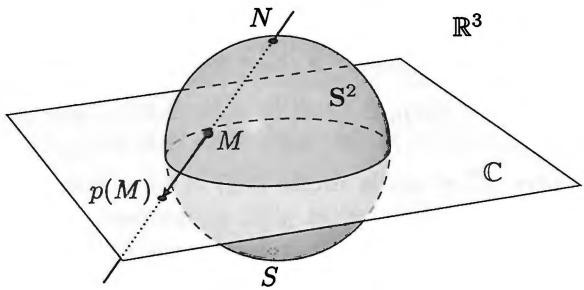


Figure 3.1. *Projection stéréographique depuis le pôle nord* $N = (0, 0, 1)$

3.1.2. Proposition. *La projection stéréographique est un homéomorphisme de la sphère euclidienne* $\mathbf{S}^2$ *sur la sphère de Riemann* $\mathbb{P}^1(\mathbb{C})$. *Elle envoie un cercle tracé sur la sphère sur un cercle de* $\mathbb{P}^1(\mathbb{C}) = \mathbb{C} \cup \{\infty\}$ *(c'est-à-dire un cercle de* $\mathbb{C}$ *ou une droite complétée du point à l'infini).*

La fin de ce paragraphe est essentiellement la preuve de cette proposition.

3.1.3. Deux homéomorphismes

Si $M = (x, y, z)$ est un point de $\mathbf{S}^2$ et $p(M) = u + iv$ est sa projection stéréographique, on a :

$$\begin{cases} u = \dfrac{x}{1-z} \\ v = \dfrac{y}{1-z} \end{cases} \quad \text{ou encore} \quad \begin{cases} x = \dfrac{2u}{u^2+v^2+1} \\ y = \dfrac{2v}{u^2+v^2+1} \\ z = \dfrac{u^2+v^2-1}{u^2+v^2+1}. \end{cases}$$

Ces formules garantissent que p est un homéomorphisme de la sphère épointée $\mathbf{S}^2 \setminus \{N\}$ sur $\mathbb{C}$. De plus, c'est, une bijection de $\mathbf{S}^2$ sur $\mathbb{P}^1(\mathbb{C})$.

Une base de voisinages de ∞ dans $\mathbb{P}^1(\mathbb{C})$ est fournie par les complémentaires des disques de centre 0 et de rayon r, avec $r \in]0, +\infty[$. Un tel disque est envoyé par p^{-1} sur la calotte sphérique $\{(x, y, z) \in \mathbf{S}^2,\ z > z_0\}$, où $z_0 = (r^2 - 1)/(r^2 + 1)$. Puisque $(r^2 - 1)/(r^2 + 1)$ décrit l'intervalle $]0, 1[$ lorsque r décrit $]0, +\infty[$, ces calottes sont une base de voisinages de N dans $\mathbf{S}^2$. En d'autres termes, p est un homéomorphisme de $\mathbf{S}^2$ sur $\mathbb{P}^1(\mathbb{C})$.

On définit de façon analogue la projection stéréographique de pôle S, que l'on note p'. Attention, on convient ici d'orienter le plan équatorial de façon compatible avec le vecteur normal $\overrightarrow{OS}$, où O est l'origine de $\mathbb{R}^3$.

Soit M un point de la sphère, $p'(M)$ est l'affixe du point P', intersection de la droite (SM) et du plan équatorial. Notons $w = p(M)$ et $w' = p'(M)$. Comme les points P et P' appartiennent au plan NSM, les arguments de w et w' sont opposés (les orientations sont choisies ainsi). Observons le plan NSM :

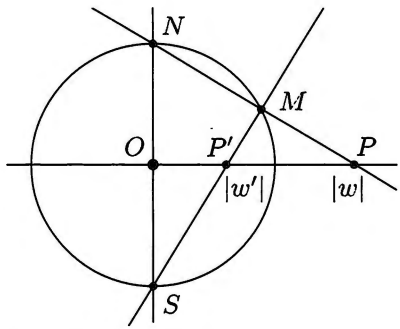


Figure 3.2. *Projection stéréographique dans le plan*

Par double orthogonalité, les angles $\widehat{NSM}$ et $\widehat{OPN}$ sont égaux : leur tangente commune vaut $|w'|/1 = 1/|w|$, d'où : $|ww'| = 1$. En définitive :

$$w' = \frac{1}{w}.$$

Ainsi, les projections stéréographiques p et p' sont compatibles avec les applications f_1 et f_2 du premier paragraphe, au sens où la restriction de p à $\mathbf{S}^2 \setminus \{N\}$ s'identifie avec f_1 et la restriction de p' à $\mathbf{S}^2 \setminus \{S\}$ s'identifie avec f_2.

3.1.4. Aparté : cocyclicité et alignement

Le critère de cocyclicité-alignement suivant est bien connu[7] (par exemple : le livre [4], corollaire III.1.20 et proposition VI.7.5). On en donne une autre preuve — plus calculatoire — que celle de [4], dans l'exercice C.4.

3.1.5. Théorème. *Quatre points du plan complexe sont cocycliques ou alignés si et seulement si leur birapport, calculé dans $\mathbb{P}^1(\mathbb{C})$, est réel.*

Cette propriété incite à appeler *cercle* de $\mathbb{P}^1(\mathbb{C})$ un cercle de $\mathbb{C} \subset \mathbb{P}^1(\mathbb{C})$ ou la réunion d'une droite de $\mathbb{C}$ et du point à l'infini.

[7] Sans dire le mot « birapport », c'est un classique de terminale S, de CAPES et même de l'Agrégation dans la thématique « Utilisation des complexes en géométrie ».

3.1.6. Corollaire. *Une homographie de $\mathbb{P}^1(\mathbb{C})$ envoie un cercle sur un cercle.*

3.1.7. Remarque. On a en fait le théorème de reconstruction suivant ([4], théorème VI.7.11) : une bijection de $\mathbb{P}^1(\mathbb{C})$ dans $\mathbb{P}^1(\mathbb{C})$ qui préserve la famille des cercles est une homographie ou la composée d'une homographie et de la conjugaison complexe. Ces bijections forment le *groupe circulaire* $\mathrm{PGL}_2(\mathbb{C}) \rtimes \mathbb{Z}/2\mathbb{Z}$.

3.1.8. Préservation des cercles par projection stéréographique

Cette méthode nous a été indiquée par Serge Parmentier. On en trouvera une autre, tirée du film *Dimensions* [2], dans l'exercice C.9.

Soit n un entier naturel non nul. On considère l'espace affine euclidien $\mathbb{R}^n$. On a surtout en vue les cas $n = 2$ et $n = 3$, mais le cas $n = 4$ a été utile pour représenter la fibration de Hopf (VII-B.2). D'autre part, rappelons qu'une inversion de $\mathbb{R}^n$ de pôle N et de rapport $k > 0$ est l'application de $\mathbb{R}^n \setminus \{N\}$ dans $\mathbb{R}^n \setminus \{N\}$ qui, au point M, associe le point M' de la demi-droite $[NM)$ d'origine N tel que :

$$NM \cdot NM' = k.$$

3.1.9. Lemme. *L'image d'une hypersphère ou d'un hyperplan par une inversion est une hypersphère ou un hyperplan.*

Démonstration. Quitte à changer de repère, on peut supposer que N est l'origine de $\mathbb{R}^n$. Quitte à composer par une homothétie, on peut supposer que k vaut 1. De la sorte, l'inversion est une involution. Les coordonnées de l'image M' d'un point $M = (x_i)_{1 \leqslant i \leqslant n}$ (distinct de N) sont :

$$x'_i = \frac{x_i}{\sum_{j=1}^{n} x_j^2} \quad (1 \leqslant i \leqslant n) \quad \text{et} \quad \sum_{i=1}^{n} x_i^2 \sum_{j=1}^{n} {x'_j}^2 = 1.$$

Une hypersphère possède une équation de la forme

$$a \sum_{i=1}^{n} x_i^2 - 2 \sum_{i=1}^{n} b_i x_i + d = 0,$$

où $a = 1$, $(b_i)_{1 \leqslant i \leqslant n}$ est le centre et d est lié au rayon r par : $r^2 = \sum_i b_i^2 - d$. En divisant par $\sum x_i^2$, on voit que cette équation équivaut à :

$$a - 2 \sum_{i=1}^{n} b_i x'_i + d \sum_{i=1}^{n} {x'_i}^2 = 0,$$

ce qui prouve que l'image de l'hypersphère est contenue dans une hypersphère (si $d \neq 0$) ou un hyperplan (si $d = 0$). En reprenant ce qui précède avec $a = 0$, on montre que l'image d'un hyperplan est contenue dans une hypersphère ou dans un hyperplan.

Grâce à l'involutivité, on conclut que l'image d'une hypersphère ou d'un hyperplan est une hypersphère ou un hyperplan (privé de N si nécessaire).□

3.1.10. Remarque. Lorsque $n = 2$, les calculs purement cartésiens ci-dessus sont indépendants de toute considération sur le birapport. Grâce à ce lemme, on peut montrer à nouveau qu'une homographie envoie un cercle ou une droite sur un cercle ou une droite, et retrouver le critère de cocyclicité-alignement précédent. Voir l'exercice C.6.

3.1.11. Lemme. *Dans $\mathbb{R}^3$, la projection stéréographique de la sphère unité $\mathbf{S}^2$ sur l'équateur coïncide avec la restriction de l'inversion de pôle N et de rapport 2.*

Le lemme exprime que l'image de $\mathbf{S}^2 \setminus \{N\}$ par l'inversion est le plan équatorial $\{z = 0\}$.

Démonstration. Soit $M = (x, y, z)$ un point de $\mathbf{S}^2$ distinct de N et $M' = (u, v, 0)$ son image par la projection stéréographique. Il suffit de montrer que $NM \cdot NM' = 2$. Pour cela, on calcule :

$$(NM \cdot NM')^2 = \Big(\frac{x^2+y^2}{(1-z)^2} + (-1)^2\Big)\big(x^2+y^2+(z-1)^2\big) = \frac{(2-2z)^2}{(1-z)^2} = 4.$$

□

3.1.12. Corollaire. *La projection stéréographique de la sphère unité $\mathbf{S}^2$ sur l'équateur envoie un cercle tracé sur la sphère sur un cercle ou une droite.*

Démonstration. Par le lemme 3.1.11, l'image du cercle par projection stéréographique coïncide avec son image par l'inversion. Or, tout cercle de $\mathbf{S}^2$ est l'intersection de $\mathbf{S}^2$ et d'un plan. Par le lemme 3.1.9, le cercle est donc l'intersection du plan équatorial et de l'image du plan, qui est une sphère ou un plan : on obtient respectivement un cercle ou une droite. □

3.1.13. Remarque. On voit bien comment définir, pour tout entier n, la projection stéréographique p de la sphère unité $\mathbf{S}^{n-1}$ de $\mathbb{R}^n$ privé du point $N = (0, \ldots, 0, 1)$ sur l'hyperplan H d'équation $x_n = 0$: l'image d'un point M est l'intersection de la droite (NM) avec l'hyperplan H. On prolonge p en posant $p(N) = \infty$, d'où un homéomorphisme entre $\mathbf{S}^{n-1}$ et le compactifié d'Alexandroff $H \cup \{\infty\}$ de $H \simeq \mathbb{A}^{n-1}$.

Dans ce nouveau contexte, le lemme 3.1.11 reste vrai (même preuve, un peu plus de coordonnées), de même que le corollaire 3.1.12 (un cercle peut toujours être construit comme une intersection d'hyperplans et de sphères). Ainsi, la projection stéréographique envoie une droite ou un cercle sur une droite ou un cercle.

3.2. Configuration de Miquel

Changement de thème pour annoncer une configuration très symétrique qui sera étudiée en détail dans le second volume, où l'on verra en plissant les yeux surgir un hypercube de cette configuration, beau développement du cube apparaissant dans la formule des six birapports.

Théorème. [Miquel] *Dans la sphère de Riemann, soit quatre cercles concourant en un point, en position générale pour cette contrainte. Pour tout choix de trois de ces quatre cercles, on considère le cercle circonscrit au triangle formé par les intersections des trois cercles.*

Alors, les quatre cercles ainsi construits sont concourants.

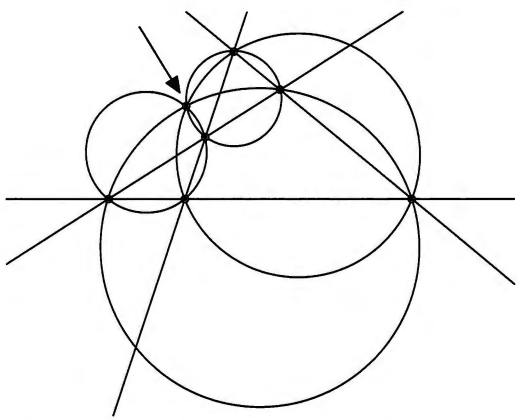


Figure 3.3. *Configuration de Miquel : quatre cercles concourant à l'infini, quatre nouveaux cercles et le point de Miquel sous la flèche*

Faisons le compte : nous avons quatre cercles au début, nous en construisons quatre nouveaux ; nous avons un point de concours initial, six points d'intersection de deux cercles pris deux à deux, et un point de concours des quatre nouveaux cercles. Au bilan, huit cercles et huit points d'intersection notables ; quatre points sur chaque cercle, quatre cercles passant par chaque point. Dans la terminologie de Hilbert, c'est une configuration de type $8_4 8_4$.

Cette configuration est extrêmement symétrique : si l'on oublie la numérotation initiale des cercles, on est bien incapable de la retrouver !

4. Birapport et configurations finies

Dans cette ébauche de paragraphe, on prend pour $\mathbb{K}$ un corps fini $\mathbb{F}_q$ de caractéristique p. La droite projective contient alors $q+1$ points. L'action de $\mathrm{PGL}_2(\mathbb{F}_q)$ sur $\mathbb{P}^1(\mathbb{F}_q)$ induit un morphisme $\mathrm{PGL}_2(\mathbb{F}_q) \to \mathfrak{S}_{q+1}$. Comme l'action est fidèle par la proposition II-4.3.6, le morphisme est injectif. Pour les petites valeurs de q, on obtient par cardinalité (rappelons que $\mathrm{PGL}_2(\mathbb{F}_q)$ contient $q(q^2-1)$ éléments ; que $\mathrm{PSL}_2(\mathbb{F}_q)$ contient $q(q^2-1)/2$ éléments si $p>2$ et $q(q^2-1)$ si $p=2$; que $\mathfrak{S}_n$ contient pour unique sous-groupe d'indice 2 le groupe alterné) :

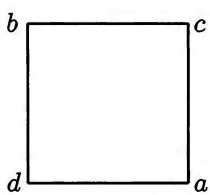
- pour $q=2$, un isomorphisme $\mathrm{PGL}_2(\mathbb{F}_2) = \mathrm{PSL}_2(\mathbb{F}_2) \simeq \mathfrak{S}_3$;
- pour $q=3$, un isomorphisme $\mathrm{PGL}_2(\mathbb{F}_3) \simeq \mathfrak{S}_4$, et, par restriction : $\mathrm{PSL}_2(\mathbb{F}_3) \simeq \mathfrak{A}_4$;
- pour $q=4$, un isomorphisme $\mathrm{PGL}_2(\mathbb{F}_4) = \mathrm{PSL}_2(\mathbb{F}_4) \simeq \mathfrak{S}_5$.

On verra au second tome une application du birapport à des configurations plus compliquées lorsque $\mathbb{K} = \mathbb{F}_5$ et, surtout, $\mathbb{K} = \mathbb{F}_7$.

A. Complément mnémotechnique sur la formule des six birapports

On reprend la situation du paragraphe 1.3 pour chercher un moyen de reconstruire la formule sans la connaître.

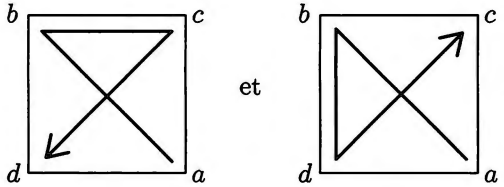
Lecture d'une face



Dans le birapport, on met côte à côte les lettres diagonalement opposées. C'est raisonnable, si l'on revient à la définition ou si l'on pense à un déterminant. Mais, compte tenu de la relation

$$[a,b,d,c] = [a,b,c,d]^{-1},$$

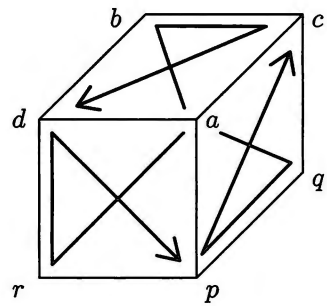
on voit qu'il importe de choisir entre



En fait, orienter une face n'a aucune importance : si une telle relation a lieu, elle a lieu avec les inverses ! Ce qui compte, c'est la compatibilité dans les ordres de lecture des faces.

Compatibilité entre faces par les sommets

Voici une première recette, où l'on tourne autour des sommets : on fixe le sens de lecture d'une face, puis l'on répercute ce choix par des rotations d'angle $2\pi/3$.

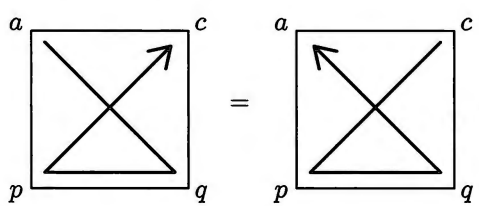


La figure précédente donne les trois premiers facteurs,

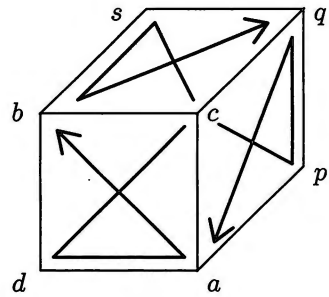
$$[a,b,c,d]\times[a,r,d,p]\times[a,q,p,c],$$

dans la formule des six birapports.

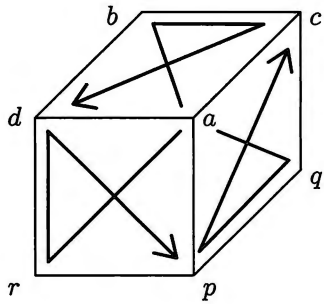
On veut maintenant changer de sommet de référence et passer à c. La relation $[a,q,p,c]=[c,p,q,a]$ se lit graphiquement :

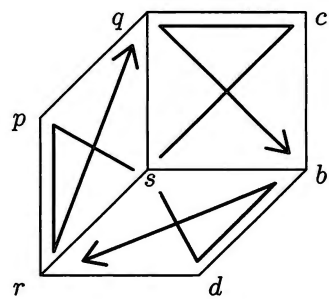


On a de même $[a,b,c,d]=[c,d,a,b]$, ce qui en faisant tourner le cube nous donne l'ordre de lecture du troisième facteur contenant c, $[c,s,b,q]$:



On veut à présent prendre s pour sommet de référence. L'orientation de la face contenant c, s, b, q est forcée par : $[c, s, b, q] = [s, c, q, b]$. On fait faire au cube initial un demi-tour d'axe (rc) :





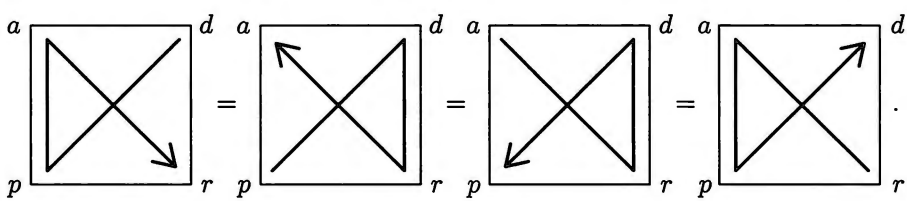
L'orientation autour de s est donc forcée, de proche en proche, par celle que l'on a choisie autour de a. En lisant les six faces du cube de cette façon, on obtient la formule des six birapports :

$$[a, b, c, d] \times [a, r, d, p] \times [a, q, p, c] \times [s, c, q, b] \times [s, p, r, q] \times [s, d, b, r] = 1.$$

Compatibilité entre faces par les arêtes

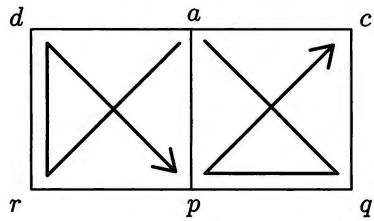
Revenons aux différentes lectures possibles de la face $[a, r, d, p]$. On lit les relations :

$$[a, r, d, p] = [r, a, p, d] = [d, p, a, r] = [p, d, r, a]$$

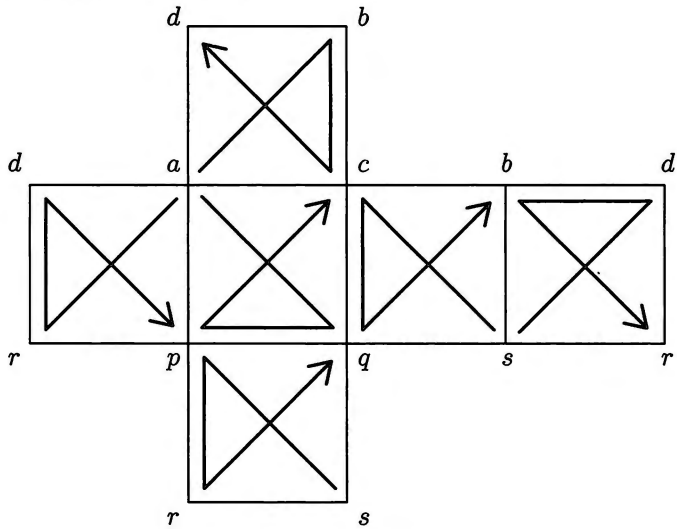


On constate que la seule chose qui compte, c'est la direction de la « barre » liant les deux diagonales : elle est horizontale. Si l'on regarde à présent deux

faces contiguës, on voit qu'il faut prendre :



D'où la règle : si l'on déplie deux faces, les « barres » doivent être l'une horizontale, l'autre verticale.



Il n'y a plus qu'à démontrer la formule des six birapports, ce qui n'est qu'une question de patience...

B. Géométrie projective dans une coquille de noix

B.0.1. Droite projective

Soit V un plan vectoriel sur $\mathbb{K}$. On appelle droite projective associée à V et l'on note $\mathbb{P}(V)$ l'ensemble des droites vectorielles de V. Le choix d'une base $\mathbf{e}$ de V donne un isomorphisme linéaire de V sur $\mathbb{K}^2$, ce qui induit une bijection (non canonique) de la droite projective $c_{\mathbf{e}} : \mathbb{P}(V) \to \mathbb{P}^1(\mathbb{K})$.

Un repère projectif de $\mathbb{P}(V)$ est un triplet de points distincts. La bijection $c_{\mathbf{e}}$ l'envoie sur un repère projectif (z_1, z_2, z_3) de $\mathbb{P}^1(\mathbb{K})$. La coordonnée d'un

point de $\mathbb{P}(V)$ dans ce repère est sa coordonnée dans le repère (z_1, z_2, z_3) de $\mathbb{P}^1(\mathbb{K})$. Si $\mathbf{e}'$ est une autre base, la composée $h_{\mathbf{e},\mathbf{e}'} = c_{\mathbf{e}'} \circ c_{\mathbf{e}}^{-1}$ provient d'un isomorphisme linéaire de $\mathbb{K}^2$ (changement de base !), donc c'est une homographie. Comme un changement de repère dans $\mathbb{P}^1(\mathbb{K})$ est une homographie, un changement de coordonnées dans $\mathbb{P}(V)$ en est une aussi.

Si $\mathbb{K}$ est $\mathbb{R}$ ou $\mathbb{C}$, on définit une topologie sur $\mathbb{P}(V)$ par transport de structure grâce à $c_{\mathbf{e}}$. Cette topologie ne dépend pas de la base choisie, car les applications $h_{\mathbf{e},\mathbf{e}'}$ (où $\mathbf{e}$ et $\mathbf{e}'$ sont deux bases de V) sont des homéomorphismes.

B.0.2. Plan projectif

Par définition, un plan projectif sur un corps $\mathbb{K}$ est l'ensemble $\mathbb{P}(V)$ des droites d'un espace vectoriel V de dimension 3. Cette définition fournit une surjection naturelle $\pi : V \setminus \{0\} \to \mathbb{P}(V)$, $v \mapsto [v] = \mathbb{K}v$, dont les fibres sont les droites. *Via* le choix d'une base, il est naturellement en bijection avec la grassmannienne $\mathbb{P}^2(\mathbb{K}) = \mathbb{P}(\mathbb{K}^3)$ des droites de $\mathbb{K}^3$, que l'on a déjà rencontrée.

Une droite projective est l'image dans le plan projectif d'un plan vectoriel (épointé, c'est-à-dire privé de l'origine) par π. C'est le noyau d'une forme linéaire ℓ appartenant au dual V^* de V. Il est bien connu que $\operatorname{Ker} \ell = \operatorname{Ker} \ell'$ si et seulement si $\ell' = \alpha\ell$ pour $\alpha \in \mathbb{K}^*$ convenable : aussi, la correspondance $\ell \mapsto \operatorname{Ker} \ell$ de $V^* \setminus \{0\}$ vers l'ensemble des droites induit une bijection de $\mathbb{P}(V^*)$ vers l'ensemble des droites de $\mathbb{P}(V)$.

Une application h du plan projectif dans lui-même est appelée homographie s'il existe une bijection linéaire ϕ de V dans lui-même telle que pour tout point $p = [v]$ du plan projectif, $h(p) = [\phi(v)]$.

Comme l'intersection de deux plans vectoriels est une droite, deux droites du plan projectif se coupent en un point. Les propriétés d'incidence sont donc bien plus simples dans le plan projectif. Précisons cependant les liens qui existent entre les deux.

Dans l'espace vectoriel $\mathbb{K}^3$, où l'on note les coordonnées x, y et z, le plan H d'équation $z = 1$ est un plan affine. Toute droite vectorielle coupe H en un point au plus, si bien que l'application $\pi : H \to \mathbb{P}(\mathbb{K}^3)$ est injective. De plus, l'image est ouverte, car le complémentaire de $\pi(H)$ est la droite projective qui est l'image du plan (épointé) d'équation $z = 0$, qui est fermé dans $\mathbb{K}^3$. Pour un plan affine quelconque $\mathscr{P}$, on choisit un repère affine pour l'identifier à H. Deux choix différents conduiront à des plongements de $\mathscr{P}$ dans $\mathbb{P}^2(\mathbb{K})$ identiques à homographie près.

Un des intérêts du plan projectif réel ou complexe est d'être compact, comme toute grassmannienne (proposition II-4.3.11 et exercice II-F.20), d'où la possibilité de le faire entrer dans une coquille de noix.

C. Exercices

Distance de Hilbert

C.1. Exercice. *Soit K l'intérieur d'un carré, d'une ellipse... Dessiner les cercles centrés en divers points de K pour la distance de Hilbert.*

Réponse en ligne : voir [21].

C.2. Exercice (Inégalité triangulaire pour la distance de Hilbert). *On s'intéresse aux cas d'égalité dans l'inégalité triangulaire. Soit x, y et z trois points d'un ouvert convexe K ne contenant pas de droite.*

1. *Vérifier que si le point z appartient au segment $[xy]$, on a alors : $d(x,z)+d(z,y)=d(x,y)$.*
2. *On suppose que K est « strictement convexe », c'est-à-dire que si deux points p et q appartiennent au bord de K, la droite (pq) ne coupe le bord de K qu'aux points p et q.*

   On suppose que x, y et z ne sont pas alignés, ce qui permet de définir x' et y' comme dans la preuve de la proposition 2.3.1. Montrer que x' et y' appartiennent à l'intérieur du convexe K et en déduire que l'on a : $d(x,z)+d(z,y)<d(x,y)$.
3. *On prend pour K le carré défini par :*

$$K=\big\{(a,b)\in\mathbb{R}^2,\ \max(|a|,|b|)<1\big\}.$$

   Montrer que si (xz) et (zy) coupent les bords verticaux du carré, par exemple pour $x=(-3/4,0)$, $y=(3/4,0)$ et $z=(0,1/10)$, on a : $d(x,z)+d(z,y)=d(x,y)$.

Cocyclicité, cercles, homographies, etc. : méthodes élémentaires

C.3. Exercice (Conjugaison dans $\mathrm{PSL}_2(\mathbb{C})$, dynamique des homographies)

On dira que deux matrices de $\mathrm{PSL}_2(\mathbb{C})$ sont conjuguées si l'on peut les relever en deux matrices conjuguées dans $\mathrm{GL}_2(\mathbb{C})$.

1. *Montrer qu'une matrice de $\mathrm{PSL}_2(\mathbb{C})$ est conjuguée*

$$\text{ou bien à } \begin{pmatrix}1 & t\\ 0 & 1\end{pmatrix},\ t\in\mathbb{C},\ \text{ ou bien à } \begin{pmatrix}\lambda & 0\\ 0 & \lambda^{-1}\end{pmatrix},\ \lambda\in\mathbb{C}^*.$$

2. *On appelle suite homographique une suite de* $\mathbb{P}^1(\mathbb{C})$ *définie par récurrence par*

$$u_{n+1} = \frac{au_n + b}{cu_n + d}, \quad \text{avec} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{C}).$$

Montrer qu'après un changement de variables, on peut transformer une suite homographique soit en une suite arithmétique soit en une suite géométrique.

C.4. Exercice (Critère de cocyclicité-alignement). *On considère* $\mathbb{C}$ *comme un plan affine euclidien. Soit* a, b, c *et* d *quatre complexes distincts.*

1. *On suppose que* a, b, c *et* d *appartiennent à un même cercle, dont on note* ω *le centre et* r *le rayon. On fixe un réel* α *tel que* $a = \omega + r\exp(i\alpha)$ *et l'on choisit de même* β, γ *et* δ. *Calculer le birapport* $[a, b, c, d]$ *en fonction des réels* α, β, γ *et* δ *et constater qu'il est réel. Traiter de même le cas où les points sont alignés.*
2. *On suppose que le birapport* $[a, b, c, d]$ *est réel.*
   (a) *On suppose que* a, b *et* c *appartiennent à une droite* Δ. *Montrer que* d *appartient aussi à* Δ.
   (b) *On pose, pour* z *complexe différent de* -1 : $h(z) = 2/(z+1)$. *Montrer que l'image du cercle unité (privé du point* -1*) par* h *est la droite* Δ *d'équation :* $\mathrm{Re}(z) = 1$ *et réciproquement.*
   (c) *On suppose que* a, b *et* c *appartiennent au cercle unité et sont distincts de* -1 *et que* $d \neq -1$. *Montrer que* $h(d)$ *appartient à la droite* Δ *et en déduire que* d *appartient au cercle unité.*
   (d) *On suppose que* a, b *et* c *ne sont pas alignés. Montrer que* d *appartient au cercle circonscrit au triangle* abc.

C.5. Exercice. *Soit* A *et* B *deux points d'un plan affine euclidien.*

1. *Soit* k *un réel strictement positif différent de* 1. *Soit* G *et* G' *les barycentres de* $\{(A,1),(B,k)\}$ *et de* $\{(A,1),(B,-k)\}$. *Montrer que l'ensemble des points* M *distincts de* A *et* B *tels que* $MA/MB = k$ *est le cercle de diamètre* $[GG']$. *Décrire les ensembles des points* M *tels que* $MA/MB = 0$, $MA/MB = 1$, $MA/MB = \infty$.
2. *Soit* α *un réel.*
   (a) *À l'aide du théorème de l'angle inscrit, montrer que les lignes de niveau*

$$\left\{ z \in \mathbb{C} \setminus \{a, b\},\ \arg \frac{z-a}{z-b} \equiv \alpha\ [\pi] \right\}$$

sont des arcs de cercles privés de a *et* b.

(b) Montrer proprement que les lignes de niveau modulo 2π sont les arcs de cercles délimités par la droite (ab).

3. *Pour a et b complexes distincts, décrire les lignes de niveau du module et de l'argument de $f : z \mapsto (z-a)/(z-b)$.*

C.6. Exercice (Critère de cocyclicité-alignement *via* les inversions). *On utilise le lemme* 3.1.9.

1. *Calculer l'affixe de l'image d'un point par une inversion. Montrer (sans calcul ; un peu de topologie suffit) que l'image d'un cercle ou d'une droite contenant le pôle d'une inversion (resp. ne le contenant pas) est une droite (resp. est un cercle).*

2. *Soit a, b, c, d quatre complexes tels que $ad - bc \neq 0$ et $c \neq 0$. En écrivant, pour z complexe tel que $cz + d \neq 0$:*

$$\frac{az+b}{cz+d} = \frac{a}{c} - \frac{ad-bc}{c(cz+d)},$$

montrer que l'image d'un cercle ou d'une droite par une homographie est un cercle ou une droite.

3. *Soit a, b, c, d quatre nombres complexes distincts. Soit h l'homographie qui envoie a sur ∞, b sur 0 et c sur 1. Montrer que a, b, c et d sont cocycliques ou alignés si et seulement si le birapport $[a,b,c,d] = h(d)$ est réel.*

C.7. Exercice (Applications de la formule des six birapports)

- ([4], figure de couverture.) *Soit $\mathscr{C}_1$, $\mathscr{C}_2$, $\mathscr{C}_3$, $\mathscr{C}_4$ quatre cercles-droites de $\mathbb{P}^1(\mathbb{C})$. On suppose que pour $i \in \{1,2,3,4\}$, $\mathscr{C}_i$ et $\mathscr{C}_{i+1}$ se coupent en A_i et A_i' (avec $\mathscr{C}_5 = \mathscr{C}_0$) et que les A_i sont cocycliques. Montrer que les A_i' sont cocycliques.*
- (Théorème du pivot, [4], exercice III.31.) *Soit un triangle ABC et soitA', B' et C' trois points des côtés (BC), (CA) et (AB) distincts des sommets. Les cercles circonscrits aux triangles $AB'C'$, $BC'A'$ et $CB'A'$ ont un point commun.*
  Le point à l'infini de $\mathbb{P}^1(\mathbb{C})$ est le huitième sommet du cube.
- (Droites de Simson et de Steiner, [4], exercices III.32 et III.33.) *Les symétriques de l'orthocentre d'un triangle par rapport aux côtés sont alignés, de même que ses projetés orthogonaux sur les côtés.*

C.8. Exercice (Théorème de Ptolémée). *Soit, dans un plan affine euclidien, quatre points A, B, C et D. On note a, b, c et d leurs affixes dans un repère orthonormé (direct). On veut démontrer l'énoncé suivant.* Les points A, B, C et D sont cocycliques ou alignés et dans cet ordre[8] si et seulement si

$$AC \cdot BD = AB \cdot CD + AD \cdot BC.$$

On se donne donc quatre points A, B, C, D d'affixes respectifs a, b, c, d.

1. Sens direct, version 1. *Calculer le birapport $[a, b, c, d]$ en supposant que ces quatre complexes ont pour module 1. En déduire que si A, B, C et D sont sur le cercle unité dans cet ordre, l'égalité de distances a lieu. Étendre le résultat à quatre points cocycliques ou alignés.*
2. Sens direct, version 2. *En utilisant la relation $[a, b, c, d] + [a, c, b, d] = 1$, redémontrer le sens direct.*
3. Sens direct, version 3 et réciproque. *Soit ι l'inversion de pôle a et de rapport 1. On note $B' = \iota(B)$, $C' = \iota(C)$ et $D' = \iota(D)$, ainsi que b', c', d' leurs affixes.*
   (a) Préliminaire. *Calculer b' en fonction de a et b et en déduire que l'on a : $B'C' = BC/(AB \cdot AC)$.*
   (b) *On suppose que A, B, C et D sont alignés ou cocycliques et dans cet ordre. Montrer que l'on a : $B'D' = B'C' + C'D'$. En déduire l'égalité de Ptolémée.*
   Il est facile de montrer que B', C' et D' sont alignés. Pour montrer que C' appartient au segment $[BD]$, on pourra vérifier que le secteur angulaire $\widehat{BAD}$, défini comme l'intersection du demi-plan de frontière (AB) contenant D et du demi-plan de frontière (AD) contenant B, est stable par ι et contient C et C'.
   (c) *On suppose que l'on a : $AC \cdot BD = AB \cdot CD + AD \cdot BC$. Montrer que les points B', C' et D' sont alignés ; en déduire que A, B, C et D sont cocycliques ou alignés dans cet ordre.*
   Même indication que ci-dessus.

C.9. Exercice (Projection stéréographique et cercles). *Cette méthode est celle du chapitre 9 du film* Dimensions *[2]. On conserve les notations du paragraphe* 3.1.

1. *Soit N et B deux points distincts de la sphère, P_N et P_B les plans tangents à la sphère en N et B et Δ l'intersection de P_N et P_B. Soit Q un plan contenant le segment $[NB]$ et M l'intersection de Q et Δ (si elle existe). Alors, on a : $NM = BM$.*
   Utiliser la réflexion de plan le plan médiateur du segment $[NB]$.

[8] Dans le cas de points alignés, que veut-on dire précisément ?

2. *Soit $\mathscr{C}$ l'intersection de la sphère et d'un plan (qui ne contient pas le centre de la sphère). À l'aide du théorème de Pythagore, vérifier s'il en est besoin que l'intersection d'une sphère et d'un plan est un cercle. Montrer que ce cercle est également l'intersection d'un cône de révolution tangent à la sphère avec la sphère. On note S le sommet du cône.*
3. *On veut montrer que la projection p' de centre N sur le plan horizontal contenant S envoie $\mathscr{C}$ sur un cercle. Soit B un point de $\mathscr{C}$ et D son image par p'. En utilisant la première question et le plan $Q = (NBS)$, montrer que $SB = SD$ et conclure.*
   Les triangles AMB et DSB sont semblables.
4. *En déduire que la projection stéréographique habituelle d'un cercle tracé sur la sphère est un cercle. (On traitera à part les cercles contenant N et les grands cercles.)*

Conjugaison harmonique

C.10. Exercice (Divisions harmoniques dans les quadrilatères complets). *Dans un plan affine ou projectif sur le corps que l'on voudra, soit p_1, p_2, q_1, q_2 quatre points non alignés trois à trois. On note m (resp. r, s) l'intersection des droites (p_1p_2) et (q_1q_2) (resp. (p_1q_1) et (p_2q_2), (p_1q_2) et (p_2q_1)) et m' (resp. n') l'intersection de (rs) avec (p_1p_2) (resp. q_1q_2) (se reporter à la figure* C.1*).*

1. *Démontrer que le quadruplet (m, m', p_1, p_2) est une division harmonique : $[p_1, p_2, m, m'] = -1$.*
   Faute d'inspiration, on pourra toujours choisir un repère d'origine m dans lequel p_1 (resp. q_1) a pour coordonnées $(1, 0)$ (resp. $(0, 1)$).

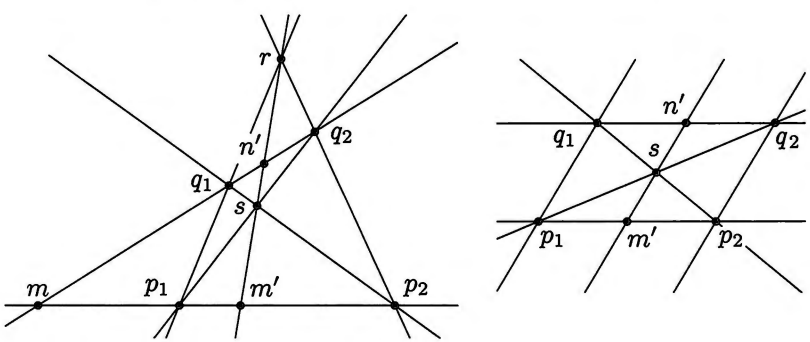


Figure C.1. *Un quadrilatère complet générique et un particulier*

2. *Un cas particulier (dans un plan projectif) est celui où la droite (mr) est la droite à l'infini. Vérifier le résultat dans ce cas et le redémontrer dans le cas général.*

C.11. Exercice (Construction du quatrième point harmonique). *À l'aide de l'exercice* C.10, *construire, à partir d'un triplet de points deux à deux distincts* (a, b, c), *un quatrième point* d *tel que* $[a, b, c, d] = -1$.

Prendre $p_1 = a$, $p_2 = b$, $m = c$, choisir un point r et une droite passant par m. Il ne reste plus qu'à reproduire la figure de l'exercice C.10 pour obtenir le point $d = m'$.

Cet exercice classique est détaillé dans le très recommandable livre [19] de Benoît Kloeckner.

C.12. Exercice (Involutions de $\mathrm{PGL}_2(\mathbb{K})$ et divisions harmoniques)

On fixe un corps $\mathbb{K}$ *de caractéristique différente de* 2.

1. *Soit* a *et* b *deux éléments distincts de* $\mathbb{P}^1(\mathbb{K})$. *Étant donné* $x \in \mathbb{P}^1(\mathbb{K}) \setminus \{a, b\}$, *montrer qu'il existe un unique* $y \in \mathbb{P}^1(\mathbb{K})$ *tel que* $[a, b, x, y] = -1$. *On note* $h_{a,b}$ *l'application qui associe* y *à* x.

   On dit que y *est le* conjugué harmonique *de* x *par rapport à* a *et* b. *Comme elle est définie en termes de birapport, qui est invariant par changement de repère projectif, la conjugaison harmonique est bien définie sur toute droite projective.*

2. *Vérifier que* $h_{a,b}$ *est une homographie, que* $h_{a,b}$ *est une involution et que* $h_{a,b}$ *fixe* a *et* b.

3. *Inversement, soit* h *une involution de* $\mathrm{PGL}_2(\mathbb{K})$. *Montrer que* h *admet exactement deux points fixes distincts* a *et* b *et que* $h = h_{a,b}$.

C.13. Exercice (Conjugaison harmonique et orthogonalité -1-). *On veut montrer que la conjugaison harmonique est une affaire d'orthogonalité. Ici,* $\mathbb{K}$ *est un corps quelconque (de caractéristique différente de* 2*). On fixe deux droites distinctes de* $\mathbb{K}^2$:

$$p_1 = \mathbb{K}\begin{pmatrix} a_1 \\ b_1 \end{pmatrix}, \quad p_2 = \mathbb{K}\begin{pmatrix} a_2 \\ b_2 \end{pmatrix}.$$

1. *Montrer qu'il existe une forme quadratique* q *sur* $\mathbb{K}^2$, *unique à scalaire près, dont le cône isotrope,* i.e. *l'ensemble des éléments isotropes est* $p_1 \cup p_2$.

   La forme $q(x, y) = 2(a_1 y - b_1 x)(a_2 y - b_2 x)$ est un candidat naturel.

2. *On identifie les droites de* $\mathbb{K}^2$ *à des points de* $\mathbb{P}^1(\mathbb{K})$, *ce qui permet de parler de leur birapport. Soit* m *et* m' *deux droites de* $\mathbb{K}^2$. *Montrer qu'elles sont orthogonales pour* q *si et seulement si* $[m, m', p_1, p_2] = -1$.

C.14. Exercice (Conjugaison harmonique et orthogonalité -2-). *On veut montrer que l'orthogonalité dans le plan affine est une affaire de conjugaison harmonique.*

On suppose un peu de familiarité avec la géométrie projective plane. On plonge le plan euclidien dans le plan projectif réel $\mathbb{P}^2(\mathbb{R})$ *et ce dernier dans le plan projectif complexe* $\mathbb{P}^2(\mathbb{C})$. *On note* $D_{\mathbb{R}}$ *la « droite à l'infini »*

$$D_{\mathbb{R}} := \{[x : y : 0],\ (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}\},$$

du plan projectif réel et l'on définit de même $D_{\mathbb{C}}$. *On interprète* $D_{\mathbb{R}}$ *comme l'ensemble des directions*

1. *Vérifier que les points* $I = [1 : i : 0]$ *et* $I' = [1 : -i : 0]$ *sont l'intersection de* $D_{\mathbb{C}}$ *avec la quadrique de* $\mathbb{P}^2(\mathbb{C})$ *d'équation* $x^2+y^2+z^2 = 0$ *(qui définit le produit scalaire canonique).*
2. *Soit* Δ *et* Δ' *deux droites de* $\mathbb{R}^2$, *et* δ *et* δ' *leurs directions : ce sont des points de* $D_{\mathbb{R}}$ *donc de* $D_{\mathbb{C}}$. *Montrer que* Δ *et* Δ' *dont orthogonales si et seulement si* δ *et* δ' *sont conjugués harmoniques par rapport à* I *et* I'.
3. *Soit* $(a, b) \in \mathbb{C}^2 \setminus \mathbb{R}^2$, *et soit* $J = [a : b : 0]$ *et* $J' = [\overline{a} : \overline{b} : 0]$ *: ce sont deux points de* $D_{\mathbb{C}}$, *qui sont « conjugués au sens complexe ». On note* $h_{J,J'} : D_{\mathbb{C}} \to D_{\mathbb{C}}$ *la conjugaison harmonique par rapport à* J *et* J' *(exercice* C.12*). Montrer que pour* $\delta \in D_{\mathbb{R}}$, *le conjugué harmonique* $\delta' = h_{J,J'}(\delta)$ *appartient à* $D_{\mathbb{R}}$. *Montrer que la correspondance* $H_{J,J'} : D_{\mathbb{R}} \to D_{\mathbb{R}}$ *qui, à la direction* δ, *associe* $h_{J,J'}(\delta)$, *est l'orthogonalité pour une forme quadratique euclidienne que l'on précisera.*
4. *Lorsque* J *et* J' *sont deux points de* $D_{\mathbb{R}}$, *étudier la correspondance* $H_{J,J'}$ *définie de façon analogue sur* $D_{\mathbb{R}} \setminus \{J, J'\}$.

En observant l'intersection du cône isotrope, i.e. *le cône des éléments isotropes, avec la droite à l'infini complexifiée, on a ainsi établi un dictionnaire qui fait correspondre* paires de points de $D_{\mathbb{C}}$ stables par conjugaison complexe *et* formes quadratiques non dégénérées sur $\mathbb{R}^2$ *: conjugaison harmonique et orthogonalité.*

Cet exercice illustre un vieux principe (qui était) bien connu, selon lequel le lieu naturel de la géométrie plane réelle est le plan projectif complexe.

C.15. Exercice (Conjugaison harmonique par rapport à une conique). *Soit* q *une forme bilinéaire non dégénérée sur* $\mathbb{K}^3$ *et* $\mathscr{C}$ *la conique associée dans* $\mathbb{P}^2(\mathbb{K}) = \mathbb{P}(\mathbb{K}^3)$. *On identifie droites (resp. plans) de* $\mathbb{K}^3$ *et points (resp. droites) de* $\mathbb{P}^2(\mathbb{K})$.

L'orthogonalité par rapport à q *établit une bijection entre droites de* $\mathbb{K}^3$ *et plans de* $\mathbb{K}^3$. *On va le traduire géométriquement dans* $\mathbb{P}^2$.

1. *Soit $m \in \mathbb{P}^2$ et d une droite contenant m qui coupe $\mathscr{C}$ en deux points p_1 et p_2. Un point de m' est dit conjugué harmonique de m par rapport à/pour $\mathscr{C}$ si $[m, m', p_1, p_2] = -1$. Avec l'exercice* C.13, *montrer que le point m' est l'intersection de l'orthogonal $m^\perp$ de m pour q et de d (vus comme des sous-espaces vectoriel de $\mathbb{K}^3$).*
2. *Soit d (resp. e) une droite issue de m. Elle coupe $\mathscr{C}$ en p_1 et p_2 (resp. q_1 et q_2). On note r (resp. s) l'intersection de (p_1q_1) et (p_2q_2) (resp. (p_1q_2) et (p_2q_1)). Enfin, on note m'' (resp. n'') l'intersection de (rs) et d (resp. e).*

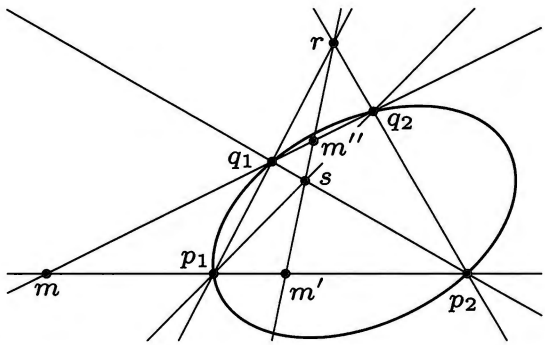


Figure C.2. *Polaire d'un point par rapport à une conique*

À l'aide de l'exercice C.10, *montrer que m'' coïncide avec le point m' de la question précédente. En déduire que la droite (rs) est la droite $m^\perp$.*
3. *Traiter les cas dégénérés (par exemple, lorsque d est tangente à $\mathscr{C}$).*

Dans le plan projectif, on appelle $m^\perp$ la polaire *de m par rapport à $\mathscr{C}$. Inversement, m s'appelle le* pôle *de m par rapport à $\mathscr{C}$.*

Deux géométries, un seul groupe

C.16. Exercice
(Transformation de Cayley et isomorphisme $SU(1,1) \simeq \mathbf{PSL_2}(\mathbb{R})$).

Dans cet exercice, $\mathscr{D}$ désigne le disque unité

$$\mathscr{D} := \{z \in \mathbb{C} \mid |z| < 1\}$$

de $\mathbb{C}$ et le mot « automorphisme » signifie une bijection biholomorphe.

1. *En utilisant le lemme de Schwarz, montrer que tout automorphisme de $\mathscr{D}$ qui fixe 0 est une rotation.* (Rappel. *Soit f une fonction holomorphe sur $\mathscr{D}$ telle que $|f(z)| \leqslant 1$ $(z \in \mathscr{D})$ et $f(0) = 0$. Alors,*

$$|f(z)| \leqslant |z| \quad (z \in \mathscr{D}), \qquad |f'(0)| \leqslant 1.$$

En particulier, s'il existe $a \in \mathscr{D} \setminus \{0\}$ tel que $|f(a)| = |a|$, il existe $c \in \mathbb{C}$ tel que $|c| = 1$ et $f(z) = cz$.)

2. *Si a et b sont deux nombres complexes tels que $|a|^2 - |b|^2 = 1$, montrer que l'homographie de $\mathbb{P}^1(\mathbb{C})$ définie par*

$$z \longmapsto \frac{az+b}{\overline{b}z+\overline{a}},$$

est un automorphisme de $\mathscr{D}$.

On pourra montrer que le cercle unité est préservé, puis utiliser une considération topologique sur la sphère de Riemann.

3. *Si w est un point de $\mathscr{D}$, trouver une homographie qui préserve $\mathscr{D}$ et envoie w sur 0. En déduire que le groupe des automorphismes de $\mathscr{D}$ est isomorphe au groupe des classes de matrices, modulo les homothéties, de la forme $\begin{pmatrix} a & b \\ \overline{b} & \overline{a} \end{pmatrix}$, telle que $|a|^2 - |b|^2 = 1$.*

4. *Reconnaître le groupe* $\mathrm{PSU}(1,1)$ *:*

$$\mathrm{PSU}(1,1) := \{A \in \mathrm{SL}_2(\mathbb{C}),\ A^*JA = J\}/\{\pm \mathrm{I}_2\}, \quad \text{où } J = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

5. *Soit $\mathscr{H}$ le demi-plan de Poincaré. Montrer que l'application g définie par*

$$g(z) = \frac{z-i}{z+i},$$

est une homographie de $\mathscr{H}$ sur $\mathscr{D}$. (On appelle g la transformation de Cayley.)

On peut remarquer tout d'abord que le cercle unité est envoyé sur la droite réelle.

6. *Montrer que le groupe des automorphismes de $\mathscr{H}$ est égal à* $\mathrm{PSL}_2(\mathbb{R})$.

Il est constitué d'homographies par ce qui précède. On montre que ce sont des homographies réelles en écrivant que l'image d'un réel est un réel.

7. *Montrer que le groupe des automorphismes de P est aussi isomorphe à* $\mathrm{PSU}(1,1)$. *En déduire que* $\mathrm{PSU}(1,1)$ *est isomorphe à* $\mathrm{PSL}_2(\mathbb{R})$.

Penser à la conjugaison.

8. *Montrer que* $\mathrm{SU}(1,1)$ *est isomorphe à* $\mathrm{SL}(2,\mathbb{R})$.

On a l'inclusion $\mathrm{SU}(1,1) \subset \mathrm{SL}(2,\mathbb{R})$ et ils ont le même quotient modulo les homothéties.

Remarque. *Conceptuellement, les choses sont assez simples : le cercle unité a pour équation $z\overline{z} - 1 = 0$, c'est donc dans $\mathbb{P}^1(\mathbb{C})$ le cône isotrope de la forme hermitienne canonique de signature $(1,1)$. L'isomorphisme exceptionnel provient juste du fait que la transformation de Cayley envoie le cercle unité sur la droite réelle.*

Chapitre XI

Trois problèmes de géométrie

On est tombées sur la tête ! On en pince pour un type qu'est mort il y a deux mille ans. *Thérèse*, Alain Cavalier, 1986.

Nous allons voir un lien entre géométrie plane et groupes de transformations sur trois exemples. On gardera en tête que ces exemples (ellipse de Steiner, théorème de Desargues, alternative de Steiner) valent plus par le procédé que pour le résultat en soi.

L'importance est donc surtout dans la méthode, et nous devons avant tout comprendre comment (et pourquoi) interviennent les groupes (groupe affine, groupe projectif réel, groupe projectif complexe). Dans les situations qui suivent, le problème géométrique est posé en terme d'invariants (droite, cercles, barycentre, alignement, parallélisme). La solution est finalement semblable dans les trois cas. Il s'agit de trouver un groupe de transformations respectant les invariants du problème (par exemple, le groupe affine respecte les droites, le barycentre, le parallélisme, mais pas les cercles). Ensuite, on montre que l'on peut se ramener à l'aide de ce groupe à une situation plus simple. Bien entendu, le lecteur pourra trouver dans la littérature (ou par lui-même) d'autres méthodes pour résoudre le problème, mais, ici, c'est l'unité de la démarche qui compte. La notion générale de repère joue un rôle central dans cette unification. On rappelle que par repère, voir section II-E, pour une géométrie, on entend un moyen d'assembler des points en une partie de sorte que le groupe définissant la dite géométrie agisse de façon simplement transitive.

Maintenant, le choix du groupe demande un minimum de subtilité. Si le groupe est trop grand, il n'aura pas assez d'invariants et donc ne conservera pas les données du problème. S'il est trop petit[1], l'orbite sur la configuration sera elle-même trop petite et l'on aura alors du mal à se retrouver dans une situation beaucoup plus agréable.

Notons pour finir que le troisième problème, sur l'alternative de Steiner, utilise les isomorphismes exceptionnels obtenus au chapitre précédent. Nous avons pensé qu'il serait intéressant de voir ces isomorphismes à l'œuvre, ou plus précisément de voir un groupe, ici $\mathrm{PSL}_2(\mathbb{C})$, régner sur deux géométries à la fois. Une nouvelle histoire, donc, dont la *tagline* s'apparenterait au "One ring to rule them all" du Seigneur des anneaux...

1. L'ellipse de Steiner

L'ellipse de Steiner donne un exemple simple de l'efficacité de l'action du groupe affine dans la résolution d'un problème de géométrie. On ramène à l'aide du groupe affine un triangle en un triangle équilatéral et ainsi le problème posé à un problème bien connu.

1.1. Énoncé du problème

1.1.1. Énoncé. Soit ABC un triangle du plan, et A', B' et C' les milieux des côtés $[BC]$, $[AC]$ et $[AB]$.

Montrer qu'il existe une ellipse tritangente au triangle aux points A', B', C'.

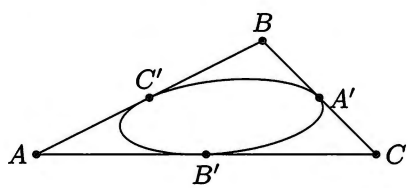


Figure 1.1. *Ellipse de Steiner*

Voici un exercice de géométrie faisant intervenir le groupe affine $\mathrm{GA}_2(\mathbb{R})$. Nous présentons ce groupe et son action naturelle sur l'espace affine.

1.2. Groupe affine et action naturelle

Le groupe affine et sa représentation linéaire a été vue en section II-5.4. Pour l'étude qui nous intéresse, nous avons l'assertion qui suit.

[1]La notion de groupe de Lie, qui possède donc une dimension en tant que variété différentielle, donne bien une idée de grandeur.

1.2.1. Proposition. *Le groupe affine* $\mathrm{GA}_2(\mathbb{R})$ *vérifie :*

$$\mathrm{GA}_2(\mathbb{R}) = \left\{ \begin{pmatrix} a & b & x_0 \\ c & d & y_0 \\ 0 & 0 & 1 \end{pmatrix} \ : \ a,b,c,d,x_0,y_0 \in \mathbb{R},\ ad-bc \neq 0 \right\} \subset \mathrm{GL}_3(\mathbb{R}).$$

Ce groupe agit naturellement sur le plan affine $\mathbb{A}^2$*, en réalisant* $\mathbb{A}^2$ *comme le plan d'équation* $z = 1$ *de* $\mathbb{K}^3$*, c'est-à-dire en assimilant le point affine* (x,y) *au vecteur* ${}^t(x\ y\ 1) \in \mathbb{K}^3$ *:*

$$\begin{pmatrix} a & b & x_0 \\ c & d & y_0 \\ 0 & 0 & 1 \end{pmatrix} \cdot (x,y) = \begin{pmatrix} a & b & x_0 \\ c & d & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = (ax+by+x_0, cx+dy+y_0).$$

1.3. Repères affines et invariants du groupe affine

Les invariants du groupe affine sont bien connus : une bijection affine g conserve :

- l'alignement (trois points alignés ont pour image trois points alignés) ;
- le parallélisme : $D \parallel D'$ et $g \in \mathrm{GA}_2 \Longrightarrow g(D) \parallel g(D')$. On vérifie que la direction de $g(D)$ est bien $\overrightarrow{g}(\overrightarrow{D})$;
- le barycentre : le barycentre de $\{(A_1,a_1),\dots,(A_n,a_n)\}$ a pour image le barycentre de $\{(g(A_1),a_1),\dots,(g(A_n),a_n)\}$.
  (Preuve : on applique à l'égalité vectorielle caractérisant le barycentre l'application linéaire $\overrightarrow{g}$ associée à g, sachant que $\overrightarrow{g}(\overrightarrow{AB}) = \overrightarrow{g(A)g(B)}$ pour tout A et B.)

1.3.1. Définition. Un repère de $\mathbb{A}^2$ est un triplet (O,I,J) de points de $\mathbb{A}^2$ non-alignés (ou tels que $(\overrightarrow{OI},\overrightarrow{OJ})$ soit une base de $\mathbb{R}^2$).

1.3.2. Proposition. *Le groupe* $\mathrm{GA}_2(\mathbb{R})$ *agit simplement et transitivement sur les repères du plan affine* $\mathbb{A}^2$*.*

Autrement dit, pour tout couple de repères (O,I,J) *et* (O',I',J')*, il existe un unique* $g \in \mathrm{GA}_2(\mathbb{R})$ *tel que* $g(O) = O'$*,* $g(I) = I'$ *et* $g(J) = J'$*.*

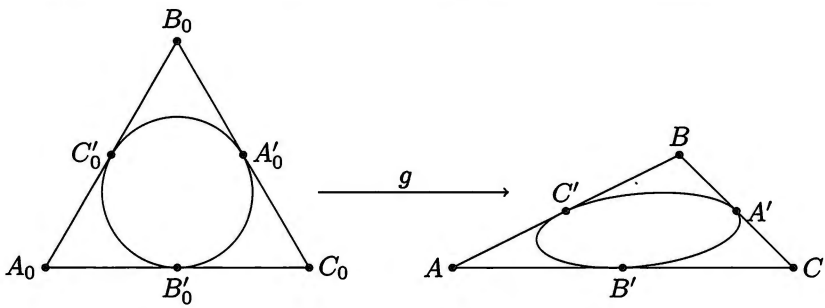
Démonstration. Nécessairement, $g(O) = O'$ et $\overrightarrow{g}$ est caractérisé de façon unique par $\overrightarrow{g}(\overrightarrow{OI}) = \overrightarrow{O'I'}$ et $\overrightarrow{g}(\overrightarrow{OJ}) = \overrightarrow{O'J'}$. □

1.3.3. Exercice. *Calculer de deux manières différentes* $|\mathrm{GA}_2(\mathbb{F}_q)|$ *: une fois, en utilisant le produit semi-direct du paragraphe* II-5.4*, une autre en comptant les repères affines.*

1.4. Solution du problème : l'ellipse de Steiner

Le problème est clair si $A_0B_0C_0$ est équilatéral. L'ellipse est dans ce cas le cercle inscrit au triangle. Le triplet (A, B, C) forme un repère affine du plan donc la proposition précédente assure l'existence de $g \in \mathrm{GA}_2(\mathbb{R})$ tel que $g(A_0) = A$, $g(B_0) = B$ et $g(C_0) = C$ (l'unicité est inutile ici). Alors :

1. les milieux sont conservés $g(A'_0) = A'$, $g(B'_0) = B'$, $g(C'_0) = C'$ par conservation du barycentre ;
2. l'image du cercle par g est une conique (puisqu'un changement de variables linéaire préserve le degré) compacte (car g est continue), c'est donc une ellipse ;
3. l'application g est différentiable, donc la tangence est conservée.



Conclusion : l'ellipse cherchée est l'image par g du cercle (unique) passant par A'_0, B'_0, C'_0.

1.4.1. Exercice. *En utilisant la même méthode, montrer que cette ellipse est unique.*

1.4.2. Remarque. On peut s'attarder sur cette figure et passer en mode « mathématiques contemplatives » (les figures sont faites pour ça !). Après un instant de méditation, on se pose la question suivante.

> Pourquoi la figure de gauche possède-t-elle une plus belle symétrie que celle de droite alors qu'elles sont dans l'orbite d'un même groupe (le groupe affine) ?

La réponse est fortement liée à la proposition VI-1.10, selon laquelle tout sous-groupe fini de $\mathrm{GL}_2(\mathbb{R})$ est conjugué à un sous-groupe de $\mathrm{O}(2)$. En effet, supposons sans perte de généralité que les deux triangles aient même centre de gravité de sorte qu'on peut les voir sous la même orbite pour $\mathrm{GL}_2(\mathbb{R})$. Alors, le stabilisateur du triangle équilatéral (clairement dans $\mathrm{O}(2)$) est le groupe fini $\mathfrak{S}_3$ (si vous ne le voyez pas directement, ce sera plus clair après le

chapitre XII), et l'on dit donc que ce triangle possède une belle symétrie. Le triangle quelconque possède un stabilisateur conjugué à $\mathfrak{S}_3$, puisque qu'il se trouve dans la même orbite, mais sa symétrie est cachée par le fait que ce groupe fini est conjugué à un sous-groupe de $\mathrm{O}(n)$ et donc un sous-groupe d'un groupe orthogonal $\mathrm{O}(q)$ pour une forme quadratique définie positive q. En bref, la proposition VI-1.10 implique que tout triangle quelconque (non aplati) est équilatéral pour une certaine distance provenant d'une forme quadratique définie positive. En ces temps d'intolérance, il est important de le signaler.

2. Le théorème de Desargues

Girard Desargues (1591 - 1661), Lyonnais.

Ici, le problème posé n'est pas une simple illustration. En effet, le théorème de Desargues est le théorème de base de la géométrie projective. Il est intéressant de savoir qu'il ne peut pas se retrouver à l'aide des axiomes de la géométrie affine. On sait aussi que dans la formalisation de la géométrie par la théorie des espaces vectoriels, « l'axiome de Desargues » est équivalent à l'associativité de la multiplication dans le corps de base. En quelque sorte, si l'on veut faire de la géométrie moderne où la propriété de Desargues ne marche pas, il faut travailler dans un espace vectoriel sur un corps non associatif (on vous laisse deviner la définition) ; les octonions en fournissent un exemple édifiant, voir exercice VII-C.5.

2.1. Énoncé du problème

2.1.1. Théorème (Desargues). *Deux triangles ABC et $A'B'C'$ du plan sont en perspective si et seulement si les points X, Y et Z sont alignés, où $X = AB \cap A'B'$, $Y = BC \cap B'C'$ et $Z = AC \cap A'C'$*

Ce problème de géométrie fait intervenir le groupe projectif $\mathrm{PGL}_3(\mathbb{R})$. Comme précédemment, nous allons présenter ce groupe (déjà défini dans un cadre plus général, voir la définition II-4.3.7), ainsi que ses invariants.

2.1.2. Remarque. Dans le cas où les triangles ABC et $A'B'C'$ sont « parallèles », les points X, Y et Z sont à l'infini, c'est-à-dire sur la droite infinie, voir remarque II-5.4.2.

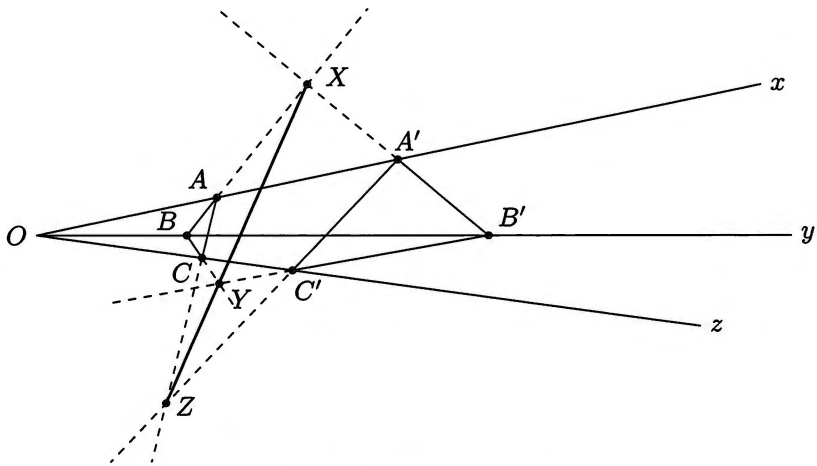


Figure 2.1. *Configuration de Desargues*

2.2. Groupe projectif réel, action naturelle et invariants

L'idée est donc d'envoyer X, Y, Z à l'infini pour obtenir la configuration 2.2, bien plus sympathique, et qui est en fait la réunion de trois configurations de Thalès. Mais, le groupe $\mathrm{GA}_2(\mathbb{R})$ conserve le parallélisme et il ne permet donc pas de se ramener à cette situation.

Nous allons donc utiliser le groupe projectif $\mathrm{PGL}_3(\mathbb{R})$ qui agit sur le plan projectif $\mathbb{P}^2(\mathbb{R})$. Rappelons la situation : un point de $\mathbb{P}^2$ est assimilé à une droite vectorielle de $\mathbb{R}^3$ et une droite de $\mathbb{P}^2$ est assimilée à un plan vectoriel de $\mathbb{R}^3$, par la projection $\pi : \mathbb{R}^3 \setminus \{0\} \longrightarrow \mathbb{P}^2$ qui associe à un vecteur non nul de l'espace la droite qu'il engendre. Le corollaire II-4.3.8 donne ce qui suit.

2.2.1. Proposition. *Le groupe projectif linéaire* $\mathrm{PGL}_3(\mathbb{R}) = \mathrm{GL}_3(\mathbb{R})/\mathbb{R}^*$ *agit de façon naturelle sur l'espace projectif* $\mathbb{P}^2(\mathbb{R}) = \mathbb{P}^1(\mathbb{R}) \sqcup \mathbb{A}^2$, *ensemble des droites de* $\mathbb{R}^3$. *Cette action est fidèle.*

Son invariant est l'alignement ; il ne conserve ni le barycentre, ni le parallélisme.

2.2.2. Remarque. Ces définitions et cette proposition peuvent être généralisées en dimension supérieure.

On montre ainsi que $\mathrm{PGL}_3(\mathbb{R})$ préserve l'alignement des points de $\mathbb{P}^2$ en remarquant par relèvement que $\mathrm{GL}_3(\mathbb{R})$ préserve la « coplanarité » des droites.

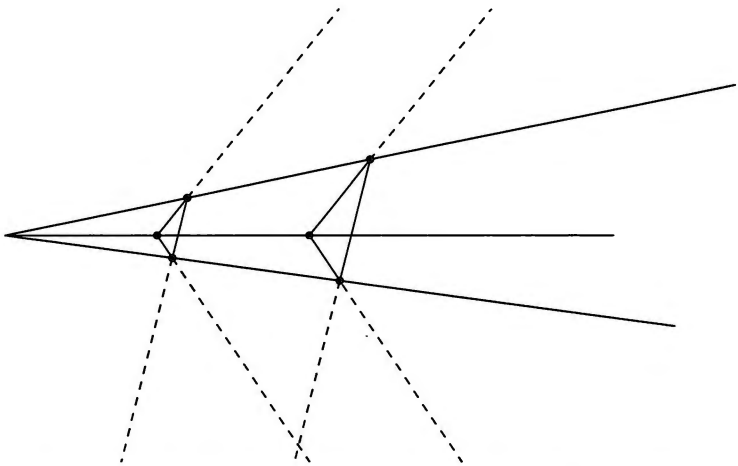


Figure 2.2. *Configuration de Desargues avec (XYZ) à l'infini*

Maintenant, nous devons plonger le plan affine sur lequel on travaille dans le plan projectif, comme nous l'avons déjà fait. Le plan affine devient une partie ouverte et dense du plan projectif pour la topologie quotient.

L'inconvénient de $\mathrm{PGL}_3(\mathbb{R})$ est qu'il ne stabilise pas $\mathbb{A}^2 \subsetneq \mathbb{P}^2 = \mathbb{A}^2 \sqcup \mathbb{P}^1$. En revanche, il est intuitivement plus gros que $\mathrm{GA}_2(\mathbb{R})$ (on montre qu'il s'agit d'un groupe de Lie de dimension $3^2 - 1 = 8$, contre la dimension $2^2 + 2 = 6$ de GA_2), donc plus « puissant » : il permet de se ramener à de plus nombreuses situations.

En l'occurrence, deux droites parallèles D et D' de $\mathbb{A}^2$ se coupent en l'infini, comme l'indique la figure 2.3.

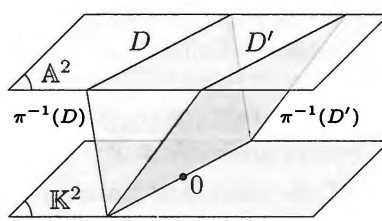


Figure 2.3.

Pour être plus explicite, ce dessin représente $\mathbb{P}^2$, qui est la réunion du plan affine correspondant au plan affine $\mathbb{A}^2$ d'équation $z = 1$ et de la droite projective correspondant aux droites du plan vectoriel $\mathbb{K}^2$ d'équation $z = 0$. Voici comment penser que deux droites parallèles se coupent en l'infini : si D et D' sont deux droites affines parallèles, alors elles définissent deux

plans $\pi^{-1}(D)$, $\pi^{-1}(D')$ dans $\mathbb{R}^3$. Et comme elles sont parallèles, ces deux plans se coupent en une droite vectorielle du plan $z = 0$, qui définit elle-même un point P de $\mathbb{P}^2$. On dit donc que les droites D et D' se coupent en le point P.

2.3. Repère projectif

Nous introduisons ici la notion de repère projectif dans la situation bien précise du plan réel. Notons que cette notion a déjà été définie en section X-1.2.9 et que l'on généralisera en toute dimension finie (et tout corps !) dans l'exercice 2.3.6.

2.3.1. Définition. Soit $(e_i)_{0\leqslant i\leqslant 2}$ une base de $\mathbb{R}^3$. On appelle repère projectif (P_0, P_1, P_2, P_3) de $\mathbb{P}^2$ un quadruplet de points de $\mathbb{P}^2$ tels que l'on ait :

(1) $(P_0, P_1, P_2) = \pi(e_0, e_1, e_2)$;

(2) $P_3 = \pi(1, 1, 1)$ dans cette base.

On vérifie que c'est équivalent à la définition suivante.

2.3.2. Définition (*bis*). On appelle repère projectif (P_0, P_1, P_2, P_3) de $\mathbb{P}^2$ un quadruplet de points de $\mathbb{P}^2$ tels que trois quelconques d'entre eux ne sont pas alignés.

En effet, la première définition implique clairement la seconde. Pour la réciproque, il suffit de partir des points P_i, $0 \leqslant i \leqslant 3$ et de choisir des relevés respectifs e'_i, $0 \leqslant i \leqslant 3$. Par hypothèse, (e'_i), $0 \leqslant i \leqslant 2$, forme une base de l'espace et de plus, e'_3 possède des coordonnées toutes non nulles (λ_i) dans cette base. Il suffit de poser $e_i = \lambda_i e'_i$, $0 \leqslant i \leqslant 2$, pour se ramener à la construction de la première définition.

2.3.3. Proposition. *Le groupe* $\mathrm{PGL}_3(\mathbb{R})$ *agit de façon simplement transitive sur l'ensemble des repères projectifs de* $\mathbb{P}^2$.

Démonstration.

En utilisant la définition *bis*, on obtient que $\mathrm{PGL}_3(\mathbb{R})$ agit bien sur l'ensemble des repères projectifs de $\mathbb{P}^2$, puisque trois droites non coplanaires ont leurs images par un élément de $\mathrm{GL}_3(\mathbb{R})$ non coplanaires.

Montrons que l'action est simplement transitive, en utilisant cette fois la première définition du repère projectif. Soit $(P_i)_{0\leqslant i\leqslant 3}$ et $(P'_i)_{0\leqslant i\leqslant 3}$ deux repères projectifs. On veut montrer qu'il existe un unique $g \in \mathrm{PGL}_3(\mathbb{R})$ tel que $g(P_i) = P'_i$ pour tout i. On veut donc trouver $\tilde{g} \in \mathrm{GL}_3(\mathbb{R})$ tel

que $\tilde{g}(e_i) = \lambda_i e'_i$, pour $0 \leqslant i \leqslant 3$ et $\lambda_i \neq 0$, et où e_i, e'_i sont des vecteurs directeurs de $\pi^{-1}(P_i), \pi^{-1}(P'_i)$, $e_3 = e_0 + e_1 + e_2$, $e'_3 = e'_0 + e'_1 + e'_2$.

Fixons λ_i, soit $\tilde{g}$ l'unique automorphisme tel que $\forall 0 \leqslant i \leqslant 2, \quad \tilde{g}(e_i) = \lambda_i e'_i$. Il reste à voir que $\tilde{g}(e_3) = \lambda_3 e'_3$. Donc

$$\lambda_0 e'_0 + \lambda_1 e'_1 + \lambda_2 e'_2 = \tilde{g}(e_0) + \tilde{g}(e_1) + \tilde{g}(e_2) = \tilde{g}(e_3) = \lambda_3 e'_3 = \lambda_3(e'_0 + e'_1 + e'_2),$$

et donc, en identifiant, les λ_i sont tous égaux. Conclusion, il existe un unique $\tilde{g} \in \mathrm{GL}_3(\mathbb{R})$ à homothétie près, donc g est unique dans $\mathrm{PGL}_3(\mathbb{R})$.□

Ce serait dommage, maintenant que la motivation est en main, de ne pas en profiter pour une généralisation bien naturelle laissée en exercice.

2.3.4. Définition. Un repère projectif $(P_0, \ldots, P_{n+1})$ de l'espace $\mathbb{P}^n(k)$ est un $(n+2)$-uplet de points tels que $n+1$ quelconques d'entre eux n'appartiennent pas à un même hyperplan.

2.3.5. Proposition. *L'action naturelle de* $\mathrm{PGL}_{n+1}(\mathbb{R})$ *sur l'ensemble des repères projectifs de* $\mathbb{P}^n$ *est simplement transitive.*

Encore une fois, un peu de combinatoire sur les corps finis constitue le meilleur test de compréhension de ces objets.

2.3.6. Exercice. *On rappelle qu'un repère projectif de* $\mathbb{P}^n$ *comprend* $n+2$ *points tels que* $n+1$ *points parmi eux n'appartiennent pas au même hyperplan. Adapter les preuves précédentes à la dimension quelconque. Montrer que* $\mathrm{PGL}_{n+1}(\mathbb{K})$ *agit de façon simplement transitive sur l'ensemble des repères projectifs. Enfin, montrer en utilisant les deux définitions du repère géométrique que :*

$$\begin{aligned} |\mathrm{PGL}_n(\mathbb{F}_q)| &= (q^n - 1) \times \cdots \times (q^2 - 1) q^{\frac{n(n-1)}{2}} \\ &= (1 + q + \cdots + q^{n-1})(q + q^2 + \cdots + q^{n-1}) \times \cdots \times q^{n-1}(q-1)^{n-1} \\ &= |\{\text{repères projectifs de } \mathbb{P}^n\}|. \end{aligned}$$

2.4. Solution du problème : théorème de Desargues

Démonstration du théorème de Desargues.

Supposons d'abord que ABC et $A'B'C'$ soient en perspective, et montrons que X, Y, Z sont alignés. On va se ramener grâce au groupe projectif réel, au cas où les triangles ont leurs côtés deux à deux « parallèles ».

D'après ce qui précède, il existe $g \in \mathrm{GL}_3(\mathbb{R})$ qui envoie X, Y sur deux points à l'infini (rappelons que l'infini est une droite projective et que deux points distincts peuvent toujours être complétés en un repère projectif).

Ainsi, les droites BC et $B'C'$ sont parallèles et de même AB et $A'B'$ sont parallèles.

Par conservation de l'alignement, montrer que X, Y, Z sont alignés revient à montrer que Z est aussi à l'infini, c'est-à-dire que les droites (AC) et $(A'C')$ sont parallèles. On reconnaît une configuration de Thalès :

$$\begin{cases} AB \parallel A'B' \\ BC \parallel B'C' \end{cases} \implies \begin{cases} \dfrac{OA}{OB} = \dfrac{OA'}{OB'} \\ \dfrac{OB}{OC} = \dfrac{OB'}{OC'} \end{cases} \implies \frac{OA}{OC} = \frac{OA'}{OC'} \implies AC \parallel A'C'.$$

La réciproque découle par *autodualité projective* de ce que l'on vient de montrer. On munit $\mathbb{R}^3$ d'une forme bilinéaire non dégénérée et l'on remarque qu'ainsi, l'orthogonalité envoie les points de $\mathbb{P}^2$ sur les droites de $\mathbb{P}^2$ et inversement.

L'orthogonalité fournit donc une bijection involutive qui inverse les inclusions et l'on vérifie que :

(1) P est un point $\iff$ $P^\perp$ est une droite,

(2) p est une droite $\iff$ $p^\perp$ est un point,

(3) $P \in p \iff p^\perp \in P^\perp$,

(4) les points P_0, P_1, P_2 sont alignés $\iff$ les droites $P_0^\perp$, $P_1^\perp$, $P_2^\perp$ sont concourantes,

(5) $p \cap q = R \iff R^\perp = (PQ)$ avec $P = p^\perp$ et $Q = q^\perp$.

En dualisant le « si » que nous venons de montrer, nous obtenons le « seulement si ». C'est l'autodualité de la configuration de Desargues. Pour le voir plus précisément...

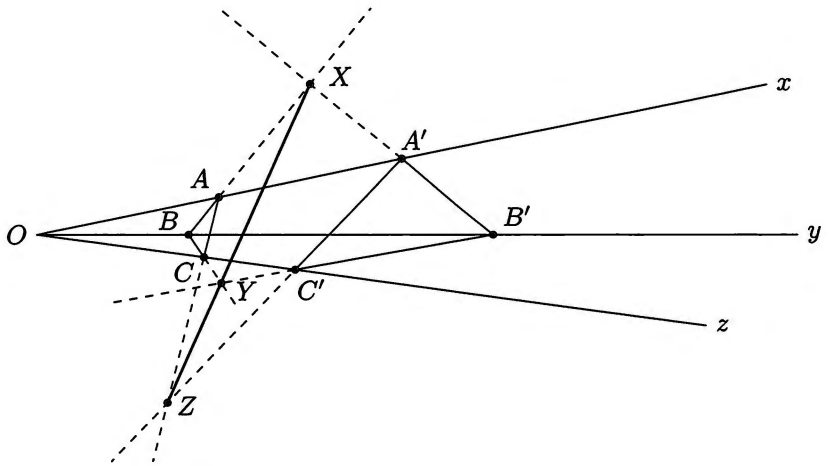
De trois droites x, y, z, on définit deux points sur chacune : $A, A' \in x$ puis $B, B' \in y$ et $C, C' \in z$.

En construisant les droites $a = BC$, $b = AC$ et $c = AB$ ainsi que $a' = B'C'$, $b' = A'C'$ et $c' = A'B'$, on obtient trois points

$$X = a \cap a', \quad Y = b \cap b', \quad Z = c \cap c',$$

que l'on sait alignés sur une droite.

Les points X, Y, Z étant alignés, les points 1 à 5 qui précèdent impliquent que les droites $X^\perp$, $Y^\perp$, $Z^\perp$ sont concourantes. Donc, par le « si » de la proposition, il vient que $x^\perp$, $y^\perp$, $z^\perp$, sont alignés (on vérifie effectivement que la construction de Desargues à partir des droites concourantes $X^\perp$, $Y^\perp$, $Z^\perp$ donne l'alignement des points $x^\perp$, $y^\perp$, $z^\perp$). Ce qui donne que $x = AA'$, $y = BB'$, $z = CC'$ sont concourantes. □



2.4.1. Remarque. Dans cette preuve, on peut, si l'on n'aime pas le choix d'un produit scalaire, utiliser l'orthogonalité « canonique » dans l'espace dual. On peut ainsi étendre l'énoncé à n'importe quel corps.

3. L'alternative de Steiner

Cette fois-ci, la solution va venir du groupe de transformations $\mathrm{PSL}_2(\mathbb{C})$ et d'un isomorphisme exceptionnel vu dans le chapitre précédent. Pour une autre preuve utilisant les inversions, on recommande l'exercice III.67 de [4].

3.1. Énoncé du problème

On considère deux cercles $\mathscr{C}$ et $\mathscr{C}'$ du plan réel tels que $\mathscr{C}$ soit à l'intérieur de $\mathscr{C}'$. On peut construire (disons, on admet qu'on peut le faire et l'on comprendra mieux pourquoi à la fin, voir 3.7) un cercle $\mathscr{C}_0$ tangent à $\mathscr{C}$ et $\mathscr{C}'$, puis un cercle $\mathscr{C}_1$ tangent à $\mathscr{C}_0, \mathscr{C}, \mathscr{C}'$, et par itération un cercle $\mathscr{C}_n$ tangent à $\mathscr{C}_{n-1}$, $\mathscr{C}$ et $\mathscr{C}'$.

Alors, soit il existe, soit il n'existe pas de $n_0 \in \mathbb{N}$, tel que $\mathscr{C}_{n_0} = \mathscr{C}_0$, *et cette alternative ne dépend que des cercles* $\mathscr{C}$ *et* $\mathscr{C}'$, *et pas du cercle* $\mathscr{C}_0$ *choisi.*

3.1.1. Remarque. Sur la figure 3.1, pour des raisons de lisibilité, les cercles $\mathscr{C}_i$ se rebouclent en un seul tour, mais ils pourraient se reboucler en plus de tours.

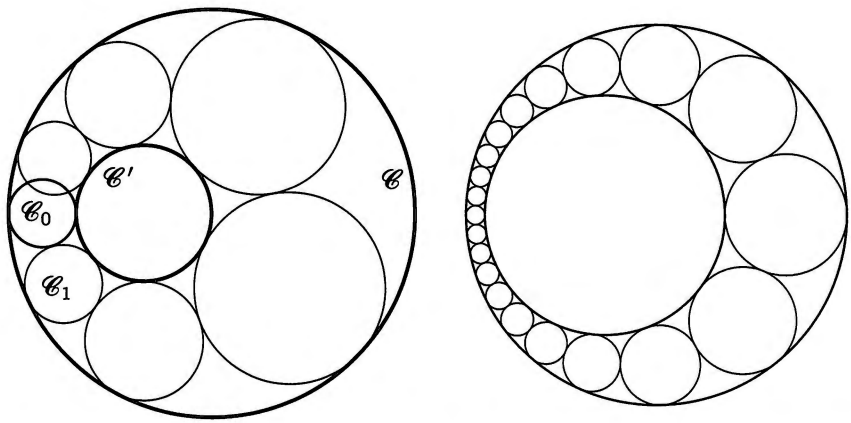


Figure 3.1. *Alternative de Steiner : $\mathscr{C}_0 \neq \mathscr{C}_7$ (à gauche) et $\mathscr{C}_0 = \mathscr{C}_{20}$ (à droite)*

3.2. Groupe $\mathrm{PGL}_2(\mathbb{C})$ et action par homographies

L'idée est de se ramener au cas où les cercles $\mathscr{C}$ et $\mathscr{C}'$ sont concentriques, cas où le problème est beaucoup plus simple. Il est clair que le groupe affine ne va pas nous y amener, car il conserve le barycentre. Il est temps de rappeler, voir section X-1.2.1, que le groupe $\mathrm{PGL}_2(\mathbb{C})$ agit sur la droite projective complexe $\mathbb{P}^1(\mathbb{C})$ et que le plan affine réel (moralement la droite complexe) est un ouvert dense de la droite projective $\mathbb{P}^1(\mathbb{C}) = \mathbb{C} \cup \{\infty\}$.

Rappelons que c'est l'identification, notée f_1, de la droite projective $\mathbb{P}^1(\mathbb{C})$ avec le compactifié $\mathbb{C} \sqcup \{\infty\}$ qui va donner naissance à l'action par homographies. À partir d'ici, pour des raisons de commodités, on pourra confondre un point de $\mathbb{P}^1(\mathbb{C})$ et son affixe. On rappelle ce qui suit, voir section X-1.2.1.

3.2.1. Proposition. *Le groupe projectif linéaire* $\mathrm{PGL}_2(\mathbb{C}) = \mathrm{GL}_2(\mathbb{C})/\mathbb{C}^*$ *agit naturellement sur la droite projective (ou sphère de Riemann)* $\mathbb{P}^1(\mathbb{C}) = \mathbb{A}^1(\mathbb{C}) \sqcup \mathbb{P}^0(\mathbb{C}) = \mathbb{A}^1(\mathbb{C}) \sqcup \{\infty\}$ *:*

$$\begin{aligned} \mathrm{PGL}_2(\mathbb{C}) \times \mathbb{P}^1(\mathbb{C}) &\longrightarrow \mathbb{P}^1(\mathbb{C}) \\ \left(\begin{pmatrix} a & c \\ b & d \end{pmatrix}, (x:y)\right) &\longmapsto (ax+cy : bx+dy). \end{aligned}$$

Ne pas confondre la droite projective complexe $\mathbb{P}^1(\mathbb{C}) = \mathbb{A}^1(\mathbb{C}) \sqcup \{\infty\}$ et le plan projectif réel $\mathbb{P}^2(\mathbb{R}) = \mathbb{A}^2(\mathbb{R}) \sqcup \mathbb{P}^1(\mathbb{R}) = \mathbb{A}^2(\mathbb{R}) \sqcup \mathbb{A}^1(\mathbb{R}) \sqcup \{\infty\}$.

Rappelons que l'on a identifié $\mathbb{P}^1(\mathbb{C}) = \mathbb{A}^1(\mathbb{C}) \sqcup \{\infty\}$ avec $\mathbb{C} \sqcup \{\infty\}$, par

transport de structure, au moyen de la bijection :

$$f_1 : \mathbb{P}^1(\mathbb{C}) \longrightarrow \mathbb{C} \sqcup \{\infty\}$$
$$(t' : t) \longmapsto \begin{cases} \frac{x}{y} & \text{si } y \neq 0 \\ \infty & \text{si } y = 0. \end{cases}$$

On peut alors reformuler l'action de $\mathrm{PGL}_2(\mathbb{C})$ sur $\mathbb{P}^1(\mathbb{C})$ en une action sur $\mathbb{C} \sqcup \{\infty\}$:

$$\mathrm{PGL}_2(\mathbb{C}) \times (\mathbb{C} \sqcup \{\infty\}) \longrightarrow \mathbb{C} \sqcup \{\infty\}$$
$$\left(\begin{pmatrix} a & c \\ b & d \end{pmatrix}, z\right) \longmapsto \begin{cases} \frac{az+c}{bz+d} & \text{si } z \neq \infty \text{ et } bz+d \neq 0 \\ \infty & \text{si } z \neq \infty \text{ et } bz+d = 0 \\ \frac{a}{b} & \text{si } z = \infty \text{ et } b \neq 0 \\ \infty & \text{si } z = \infty \text{ et } b = 0. \end{cases}$$

3.2.2. Remarque. Les règles de calcul avec l'infini ont déjà été établies en X-1.2. En tenant compte du fait que deux colonnes d'une matrice inversible ne sont pas proportionnelles, on obtient la formule plus générale

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \cdot z = \frac{az+c}{bz+d}.$$

3.3. Invariant de l'action de $\mathrm{PGL}_2(\mathbb{C})$ par homographies : birapport

Nous pouvons donc considérer que $\mathrm{PGL}_2(\mathbb{C})$ agit sur le plan réel (identifié à la droite complexe) plus un point à l'infini en prime. Les invariants de cette action ne sont ni l'alignement réel, ni le rapport $\frac{z_2 - z_1}{z_2 - z_0}$, contrairement à l'action de $\mathrm{GA}_1(\mathbb{C})$, qui conserve le rapport par Thalès. On a toutefois un invariant fondamental : le birapport (ou rapport des rapports) que l'on peut définir sur $\mathbb{C} \sqcup \{\infty\}$ par :

$$[z_0, z_1, z_2, z_3] = \frac{z_2 - z_1}{z_2 - z_0} : \frac{z_3 - z_1}{z_3 - z_0},$$

pour tout quadruplet d'éléments deux à deux distincts de $\mathbb{P}^1(\mathbb{C})$. On rappelle la proposition X-1.2.13 dans ce contexte.

3.3.1. Proposition. *Pour toute homographie g dans $\mathrm{PGL}_2(\mathbb{C})$ et tout quadruplet (z_0, z_1, z_2, z_3) de points deux à deux distincts dans $\mathbb{P}^1(\mathbb{C})$, on a :*

$$[g \cdot z_0, g \cdot z_1, g \cdot z_2, g \cdot z_3] = [z_0, z_1, z_2, z_3].$$

3.4. Repère de la droite projective complexe et cocyclicité

La notion de repère (projectif) de la droite projective a déjà été introduite en X-1.2.9 ; on ne fait ici qu'observer sa construction dans le cadre de la droite projective complexe et comment le birapport $[z_0, z_1, z_2, z_3]$ apparaît comme la coordonnée du point d'affixe z_3 dans le repère projectif formé par (z_0, z_1, z_2). On en déduira que l'ensemble des « droites ou cercles » est invariant par homographie.

3.4.1. Définition. Un repère projectif de $\mathbb{P}^1(\mathbb{C})$ est un triplet de points (z_0, z_1, z_2) deux à deux distincts. Un « bon » repère, *i.e.* une forme normale, est $(0, \infty, 1)$.

Voici donc un cas particulier de la proposition X-1.2.7.

3.4.2. Proposition. *Le groupe* $\mathrm{PGL}_2(\mathbb{C})$ *agit de façon simplement transitive sur l'ensemble des repères projectifs de* $\mathbb{P}^1(\mathbb{C})$.

3.4.3. Remarque. On voit par un calcul rapide que si (z_0, z_1, z_2) est un repère de $\mathbb{P}^1(\mathbb{C})$ et si z_3 est distinct de ces trois points, alors z_3 peut être retrouvé à partir de la donnée de $[z_0, z_1, z_2, z_3]$. Quand z_3 parcourt $\mathbb{P}^1(\mathbb{C})\backslash\{z_0, z_1, z_2\}$ plus toute valeur de $\mathbb{P}^1(\mathbb{C})\backslash\{0, \infty, 1\}$ peut être ainsi atteinte. Mieux, si le repère projectif est $(0, \infty, 1)$, alors $[0, \infty, 1, z] = z$. Il pourra être pratique d'autoriser z_3 d'être égal à un des points du repère afin d'atteindre tous les éléments de $\mathbb{P}^1(\mathbb{C})$ par le birapport. On voit alors que le birapport $[z_0, z_1, z_2, z_3]$ peut être vu comme une coordonnée du « point » z_3.

Voici (quasiment) une reformulation de la proposition X-1.2.13. Nous en donnons une preuve alternative.

3.4.4. Proposition. *Soit* (z_0, z_1, z_2, z_3), (z'_0, z'_1, z'_2, z'_3) *deux quadruplets de points distincts.*

Alors, il existe $g \in \mathrm{PGL}_2(\mathbb{C})$ *tel que* $g \cdot z_i = z'_i$ *pour tout* i *si et seulement si* $[z'_0, z'_1, z'_2, z'_3] = [z_0, z_1, z_2, z_3]$

Démonstration. On a déjà montré que l'existence de g entraîne l'égalité des birapports.

Réciproquement, si l'on fixe z'_0, z'_1 et z'_2, l'égalité $[z'_0, z'_1, z'_2, z] = [z_0, z_1, z_2, z_3]$ est une équation en z de degré 1 ou 0, donc possède une unique solution, ou aucune, ou tout complexe est solution. Mais, comme z'_3 est solution, l'ensemble des solutions n'est pas vide ; de plus, comme les z_i (resp. z'_i)

sont distincts deux à deux, on voit que l'application $z \mapsto [z'_0, z'_1, z'_2, z]$ n'est pas constante. En conclusion, $z = z_3$ est l'unique solution.

De plus, si l'on note g l'homographie telle que $g \cdot z_i = z'_i$, qui existe d'après ce qui précède, alors $[z'_0, z'_1, z'_2, g \cdot z_3] = [g \cdot z_0, g \cdot z_1, g \cdot z_2, g \cdot z_3] = [z_0, z_1, z_2, z_3]$. D'où, par l'unicité de la solution z ci-dessus : $g \cdot z_3 = z'_3$. □

On étudie maintenant l'ensemble des cercles (réels) de $\mathbb{P}^1(\mathbb{C})$. On rappelle la terminologie déjà définie en section X-3.1.4 que par cercle, on entend, cercle de l'espace euclidien $\mathbb{R}^2$ identifié à $\mathbb{C}$ ou droite réelle de $\mathbb{R}^2$ à laquelle on ajoute le point à l'infini.

On précise le corollaire X-3.1.6.

3.4.5. Corollaire. *Le groupe* $\mathrm{PGL}_2(\mathbb{C})$ *stabilise l'ensemble* $\mathscr{DC}$ *des droites et cercles du plan réel.*

L'action de $\mathrm{PGL}_2(\mathbb{C})$ *sur* $\mathscr{DC}$ *est transitive.*

Démonstration. La première assertion est le corollaire X-3.1.6, grand classique que tout bachelier scientifique devrait avoir dans sa musette. La transitivité provient directement de la transitivité de $\mathrm{PGL}_2(\mathbb{C})$ sur les repères projectifs, vu dans la proposition XI-3.4.2. □

3.5. Interprétation de l'action sur $\mathscr{DC}$ par homographie via les matrices hermitiennes à l'aide de l'isomorphisme exceptionnel $\mathbf{PSL_2}(\mathbb{C}) \simeq \mathbf{SO_0(3,1)}$

Rappelons que $\mathrm{PGL}_2(\mathbb{C}) = \mathrm{PSL}_2(\mathbb{C})$ (lemme X-2.2). Nous venons de voir que l'action de $\mathrm{PSL}_2(\mathbb{C})$ sur $\mathscr{DC}$ est transitive. Selon une méthode devenue classique dans le contexte « groupes et géométrie », lorsqu'une action est transitive, on regarde si elle est doublement transitive, triplement transitive, etc., jusqu'à avoir la n-transitivité avec n maximal. On s'intéresse alors à un invariant total pour la $(n+1)$-action. Ici, il n'y aura pas de double transitivité, mais c'est un invariant total sur l'action sur les couples qui va nous donner la preuve de l'alternative de Steiner.

Commençons par remarquer qu'un élément de $\mathscr{DC}$ (un « droitoucercle » ou cercle-droite) est donné par une équation de l'une des formes suivantes :
- soit $a(x^2 + y^2) + 2bx + 2cy + d = 0$, où a, b, c, d sont fixés dans $\mathbb{R}$, avec (x, y) dans $\mathbb{R}^2$,
- soit, de façon équivalente, $az\overline{z} + (b - ic)z + (b + ic)\overline{z} + d = 0$, avec $z \in \mathbb{C}$.

Dans ces équations, le quadruplet (a, b, c, d) est unique à scalaire près.

3.5.1. Lemme. *Soit* $(a,b,c,d) \in \mathbb{R}^4$. *L'équation*

$$a(x^2+y^2)+2bx+2cy+d=0$$

est l'équation d'un cercle de rayon non nul ou d'une droite (resp. d'un point) si et seulement si $b^2+c^2-ad>0$ *(resp.* $b^2+c^2-ad=0$*).*

Démonstration. Supposons d'abord a nul. Alors, l'équation est celle d'une droite si et seulement si $(b,c)\neq(0,0)$, c'est-à-dire : $b^2+c^2-ad>0$.

Supposons à présent que a n'est pas nul. On a, pour tout $(x,y)\in\mathbb{R}^2$:

$$a(x^2+y^2)+2bx+2cy+d=a\left[\left(x+2\frac{b}{a}\right)^2+\left(y+2\frac{c}{a}\right)^2-\frac{b^2+c^2-ad}{a^2}\right].$$

Si $b^2+c^2-ad\geqslant 0$, on reconnaît l'équation du cercle de centre $(-b/a,-c/a)$ et de rayon $(b^2+c^2-ad)^{1/2}/a$; sinon, il n'y a pas de solution (x,y) réelle.□

On va montrer la double transitivité de l'action de $\mathrm{PSL}_2(\mathbb{C})$ sur $\mathscr{D}\mathscr{C}$ grâce à un dictionnaire entre cercles-droites et matrices hermitiennes, 2×2 et de signature $(1,1)$ (à homothétie près).

On rappelle que l'espace des matrices hermitiennes 2×2 est :

$$\mathscr{H}_2(\mathbb{C})=\left\{\begin{pmatrix} a & b+ic\\ b-ic & d\end{pmatrix}\in\mathscr{M}_2(\mathbb{C})\ :\ a,b,c,d\in\mathbb{R}\right\}.$$

Dans cet espace vectoriel *réel*, on considère l'ensemble $\mathscr{H}_2^{(1,1)}(\mathbb{C})$ des matrices hermitiennes 2×2 de signature $(1,1)$. Pourquoi ? Parce que...

3.5.2. Remarque. Dire qu'une matrice hermitienne 2×2 a pour signature $(1,1)$, c'est dire que ses deux valeurs propres (réelles) sont de signe contraire, ou encore que son déterminant est strictement négatif. Or, pour une matrice $H\in\mathscr{H}_2(\mathbb{C})$, écrivons comme ci-dessus :

$$H=\begin{pmatrix} a & b+ic\\ b-ic & d\end{pmatrix}:\quad \text{alors, } -\det H=b^2+c^2-ad.$$

Rappelons que $\mathrm{SL}_2(\mathbb{C})$ agit sur $\mathscr{H}_2(\mathbb{C})$ par l'action de congruence (inverse) :

$$g\cdot H=g^{*-1}Hg^{-1}\quad\big(g\in\mathrm{SL}_2(\mathbb{C}),\ H\in\mathscr{H}_2(\mathbb{C})\big).$$

Mieux ! L'action passe au quotient et stabilise $\mathscr{H}_2^{(1,1)}$.

3.5.3. Lemme. *L'action de* $\mathrm{SL}_2(\mathbb{C})$ *sur* $\mathscr{H}_2(\mathbb{C})$ *se factorise en une action de* $\mathrm{PSL}_2(\mathbb{C})$. *L'ensemble* $\mathscr{H}_2^{(1,1)}(\mathbb{C})$ *est stable par* $\mathrm{PSL}_2(\mathbb{C})$.

Démonstration. En effet, pour $H \in \mathscr{H}_2(\mathbb{C})$ et $g \in \mathrm{SL}_2(\mathbb{C})$, on a : $(-g)\cdot H = g \cdot H$, donc c'est en réalité $\mathrm{PSL}_2(\mathbb{C})$ qui agit. De plus, on a : $\det(g \cdot H) = \det(g)^{-2}\det(H)$, d'où la stabilité des matrices de signature $(1,1)$, qui sont caractérisées par : $\det H < 0$. □

D'autre part, on considère l'ensemble $\mathscr{DC}$ sur lequel $\mathrm{PSL}_2(\mathbb{C})$ agit naturellement.

3.5.4. Proposition. *À une matrice hermitienne*

$$H = \begin{pmatrix} a & b+ic \\ b-ic & d \end{pmatrix} \in \mathscr{H}_2^{(1,1)}(\mathbb{C}),$$

on associe le cercle-droite

$$\mathscr{C}_H = \big\{(x,y) \in \mathbb{R}^2 \,:\, a(x^2+y^2) + 2bx + 2cy + d = 0\big\}.$$

Cette application induit une bijection $\beta : \mathscr{H}_2^{(1,1)}(\mathbb{C})/\mathbb{R}^* \to \mathscr{DC}$, *définie par* $\mathbb{R}^*H \mapsto \mathscr{C}_H$, *qui est compatible avec l'action de* $\mathrm{PSL}_2(\mathbb{C})$:

$$\beta(g \cdot H) = g \cdot \beta(H) \quad \big(g \in \mathrm{PSL}_2(\mathbb{C}),\ H \in \mathscr{H}_2^{(1,1)}(\mathbb{C})\big).$$

Démonstration. Pour $H \in \mathscr{H}_2^{(1,1)}(\mathbb{C})$ et $k \in \mathbb{R}^*$, on a : $\mathscr{C}_{kH} = \mathscr{C}_H$, donc β est bien définie. D'autre part, un cercle-droite est déterminé par son équation, laquelle est déterminée à scalaire près. Le lemme 3.5.1 et la remarque 3.5.2 permettent d'affirmer que β est bijective.

Montrer que β commute à l'action de $\mathrm{PSL}_2(\mathbb{C})$ peut se faire de manière brutale. La géométrie permet de le faire plus subtilement. À $H \in \mathscr{H}_2^{(1,1)}(\mathbb{C})$, on associe d'abord :

$$\Gamma_H = \left\{X = \begin{pmatrix} z \\ w \end{pmatrix} \in \mathbb{C}^2 \,:\, X^*HX = 0\right\}.$$

Dans un deuxième temps, on réalise le plan affine réel comme la droite complexe d'équation $w = 1$ dans $\mathbb{C}^2$ (pour compléter et parler de cercles de $\mathbb{P}^1(\mathbb{C})$ au lieu de cercles-droites du plan, on pourrait injecter cette droite affine dans la droite projective). On coupe Γ_H par cette droite ; on a :

$$\mathscr{C}_H = \Gamma_H \cap \big\{(z,w) \in \mathbb{C}^2 \,:\, w = 1\big\}.$$

En effet, pour $z = x + iy \in \mathbb{C}$ et $X = {}^t(z\ 1)$, on a :

$$(\overline{z}\ \ 1)\begin{pmatrix} a & b+ic \\ b-ic & d \end{pmatrix}\begin{pmatrix} z \\ 1 \end{pmatrix} = a(x^2+y^2) + 2bx + 2cy + d.$$

Mais alors, pour $\lambda \in \mathbb{C}$, on a : $(\lambda X)^*HX = |\lambda^2| X^*HX$. Par suite, Γ_H est

un cône[2] de $\mathbb{C}^2$. Deux points permettent de conclure. D'une part, l'action naturelle de $\mathrm{SL}_2(\mathbb{C})$ sur $\mathbb{C}^2$ induit une action de $\mathrm{PSL}_2(\mathbb{C})$ sur l'ensemble des cônes $\{\Gamma_H,\ H \in \mathscr{H}_2^{(1,1)}(\mathbb{C})\}$, action pour laquelle on a : $g \cdot \Gamma_H = \Gamma_{g \cdot H}$ pour tout g et tout H. En effet :

$$X \in \Gamma_{g \cdot H} \Leftrightarrow X^*(g \cdot H)X = 0 \Leftrightarrow (g^{-1}X)^* H (g^{-1}X) = 0 \Leftrightarrow X \in g \cdot \Gamma_H.$$

D'autre part, le passage de $\mathbb{C}^2$ à $\mathbb{P}^1(\mathbb{C})$ transforme l'action de $\mathrm{PSL}_2(\mathbb{C})$ sur les cônes en l'action de $\mathrm{PSL}_2(\mathbb{C})$ par homographies sur $\mathscr{D}$. Cela provient du fait que Γ_H est le cône engendré par son intersection avec la droite affine complexe $\{w = 1\}$ et que l'on a :

$$\mathbb{C}\begin{pmatrix} \dfrac{\alpha t + \beta}{\gamma t + \delta} \\ 1 \end{pmatrix} = \mathbb{C}\begin{pmatrix} \alpha t w + \beta w \\ \gamma t w + \delta w \end{pmatrix} \quad (t, w\alpha, \beta, \gamma, \delta \in \mathbb{C},\ \alpha\delta - \beta\gamma = 1).$$

Conclusion : la bijection β commute à l'action de $\mathrm{PSL}_2(\mathbb{C})$. □

3.5.5. Aparté : invariant anallagmatique

Résumons : nous avons injecté $\mathscr{D}\mathscr{C}$ dans l'espace projectif réel $\mathbb{P}\mathscr{H}_2$. De plus, cette injection est compatible avec les actions de $\mathrm{PSL}_2(\mathbb{C})$, d'une part par homographies sur $\mathscr{D}\mathscr{C}$, d'autre part, par congruence hermitienne sur $\mathbb{P}\mathscr{H}_2$. Or, comme nous l'avons vu à la proposition IX-2.3, l'action de $\mathrm{PSL}_2(\mathbb{C})$ sur $\mathscr{H}_2$ conserve la forme quadratique $q = -\det$, de signature $(3,1)$ (et même plus puisqu'elle fournit un isomorphisme exceptionnel de $\mathrm{PSL}_2(\mathbb{C})$ sur $\mathrm{SO}_0(3,1)$).

Quel est l'équivalent géométrique de cette propriété de conservation de q pour l'action ? Autrement dit, lorsqu'on agit par homographie sur un couple de cercles, quel invariant géométrique peut-on espérer ?

La réponse est la suivante : il s'agit de l'invariant anallagmatique, qui se voit géométriquement lorsque les cercles s'intersectent, et que l'on doit se contenter d'imaginer s'ils ne s'intersectent pas.

Comme les homographies sont des applications holomorphes, elles sont conformes.[3] Aussi, si deux cercles-droites se coupent, l'angle formé par les tangentes aux points de contact est invariant par homographie.

[2] C'est le *cône isotrope* de la forme hermitienne de matrice H. Cette notion naturelle est un peu moins familière que pour les formes quadratiques, car l'on n'étudie en général que les formes hermitiennes définies positives, dont le cône est trivial.

[3] Une application (différentiable) est dite *conforme* si elle préserve les angles (entre vecteurs tangents à des courbes). La différentielle d'une fonction holomorphe f en un point d'affixe z_0 est la multiplication par $f'(z_0)$, qui est une similitude du plan réel $\mathbb{C}$ et qui préserve donc les angles.

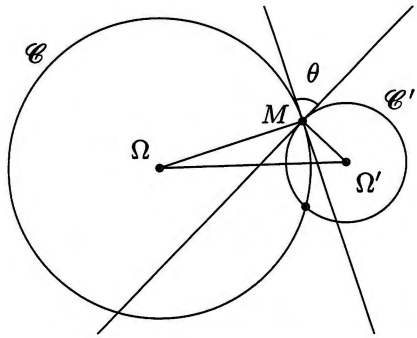


Figure 3.2. *Invariant anallagmatique de deux cercles sécants*

Soit $\mathscr{C}$ et $\mathscr{C}'$ deux cercles de centres Ω et Ω' et de rayons R et R'. On suppose qu'ils se coupent, et l'on note alors θ l'angle formé par leurs tangentes (si $\psi = \widehat{\Omega M \Omega'}$ est l'angle formé par les normales, on a $\theta = \pi - \theta$ ou peu s'en faut[4]). D'après le théorème d'Al Kashi, ce fameux théorème de Pythagore généralisé, on a la relation :

$$\Omega\Omega'^2 = R^2 + R'^2 + 2RR'\cos\theta.$$

Cela motive pour essayer d'écrire le rapport $(R^2 + R'^2 - 2\Omega\Omega')/2RR'$ en termes d'équations.

3.5.6. Lemme. *Soit φ la forme bilinéaire symétrique associée à la forme quadratique $q = -\det$ sur $\mathscr{H}_2(\mathbb{C})$. Soit*

$$H = \begin{pmatrix} a & b+ic \\ b-ic & d \end{pmatrix} \quad \text{et} \quad H' = \begin{pmatrix} a' & b'+ic' \\ b'-ic' & d' \end{pmatrix}$$

deux matrices hermitiennes de signature $(1,1)$ correspondant aux cercles $\mathscr{C}_H$ et $\mathscr{C}_{H'}$. Soit Ω et Ω' leurs centres et R et R' leurs rayons. Alors :

$$a^2R^2 = q(H), \quad a'^2R'^2 = q(H'), \quad a^2a'^2(R^2 + R'^2 - \Omega\Omega'^2)2aa'\varphi(H,H').$$

Démonstration. La preuve du lemme 3.5.1 donne :

$$a^2R^2 = b^2 +^2 -ad = -\det(H) = q(H), \quad a'^2R'^2 = q(H'),$$

et, d'autre part :

$$a^2a'^2\Omega\Omega'^2 = (ab' - a'b)^2 + (ac' - a'c)^2.$$

[4]Sur la figure, tout a l'air clair mais il y a une vraie ambiguïté en réalité : voir la remarque 3.5.9.

On trouve après calcul :

$$a^2a'^2(R^2+R'^2-\Omega\Omega'^2)=aa'(2bb'+2cc'-a'd-ad')=2aa'\varphi(H,H').$$

□

3.5.7. Corollaire. *On reprend les notations du lemme* 3.5.6. *On suppose* $aa'\neq 0$.

(i) *Les cercles* $\mathscr{C}_H$ *et* $\mathscr{C}_{H'}$ *se coupent si et seulement si*

$$|\varphi(H,H')|\leqslant\sqrt{q(H)q(H')}.$$

(ii) *Si* $\mathscr{C}_H$ *et* $\mathscr{C}_{H'}$ *se coupent et que* θ *désigne l'angle des tangentes aux points d'intersection (figure* 3.2*), alors :*

$$|\cos\theta|=\Big|\frac{R^2+R'^2-\Omega\Omega'^2}{2RR'}\Big|=\Big|\frac{\varphi(H,H')}{\sqrt{q(H)q(H')}}\Big|.$$

Démonstration. Rappelons que les deux cercles se coupent si et seulement si $|R-R'|\leqslant\Omega\Omega'\leqslant R+R'$, ce qui, avec le lemme 3.5.6, donne la première assertion. Quant à la deuxième, c'est une reformulation de la relation d'Al Kashi. □

Comme $\mathrm{PSL}_2(\mathbb{C})$ agit sur $\mathscr{H}_2(\mathbb{C})$ par des isométries de q (préservation du déterminant), le réel

$$\alpha(H,H')=\Big|\frac{\varphi(H,H')}{\sqrt{q(H)q(H')}}\Big|=\Big|\frac{R^2+R'^2-\Omega\Omega'^2}{2RR'}\Big|,$$

est invariant pour l'action de $\mathrm{PSL}_2(\mathbb{C})$ et il s'interprète géométriquement tout simplement comme $|\cos\theta|$, où θ est l'angle entre les deux cercles comme indiqué sur la figure 3.2.

3.5.8. Définition. On appelle invariant anallagmatique de deux cercles $\mathscr{C}$ et $\mathscr{C}'$ non dégénérés de centres respectifs Ω et Ω' et de rayon respectifs R et R' le réel :

$$\Big|\frac{R^2+R'^2-\Omega\Omega'^2}{2RR'}\Big|.$$

3.5.9. Remarque. Il est désagréable d'avoir des valeurs absolues. Elles sont pourtant nécessaires. Le rapport $\alpha=(R^2+R'^2-2\Omega\Omega')/2RR'$ *n'est pas* invariant par homographie : en effet, il est strictement plus grand que 1 si l'un des cercles est intérieur à l'autre et strictement plus petit que -1 si les cercles sont extérieurs l'un à l'autre. Or, une homographie peut changer la position respective de deux cercles : par exemple, $z\mapsto 1/z$ stabilise le cercle unité et envoie le cercle de diamètre $[1/5,1/3]$, qui est intérieur, sur le

cercle de diamètre $[3,5]$, qui est extérieur. Aussi, le rapport α peut changer de signe par homographie.

Ce qui est bien défini, en revanche, c'est $|\cos\theta|$. De fait, comme les tangentes sont des droites, θ est *a priori* un angle de droites ; comme deux cercles ont deux points d'intersection entre lesquels on ne sait pas choisir, et qui sont symétriques par rapport à la droite des centres, θ est un angle non orienté de droites.

Du côté des matrices, H et $-H$ définissent le même cercle, mais passer de l'une à l'autre change le signe de $\varphi(H,H')$, qui n'a donc pas de sens géométrique.

On laisse au lecteur le soin d'étendre les considérations précédentes au cas d'un cercle et une droite, puis au cas de deux droites. Voici un résumé dans un tableau de correspondance entre la géométrie des cercles-droites et l'espace quadratique des matrices hermitiennes.

cercles-droites $\mathscr{DC}$	matrices hermitiennes $\mathscr{H}_2^{(1,1)}/\mathbb{R}^*$
$\mathscr{C} = \{ax^2+2bx+2cy+d=0\}$	$H = \begin{pmatrix} a & b+ic \\ b-ic & d \end{pmatrix}$
$\mathscr{C}$ a des points	$q(H) = -\det H > 0$
$\mathscr{C}$ est un point	$q(H) = 0$
$\mathscr{C}$ est une droite	$a = 0$
$\mathscr{C} \cap \mathscr{C}' \neq \emptyset$	$\varphi(H,H')^2 \leqslant q(H)q(H')$
$\theta =$ angle d'intersection	$\arccos \dfrac{\varphi(H,H')}{\sqrt{q(H)q(H')}}$
cercles orthogonaux	invariant anallagmatique nul

Figure 3.3. *Dictionnaire entre cercles-droites et matrices hermitiennes*

3.5.10. Exercice. *Montrer que les cercles $\mathscr{C}$ et $\mathscr{C}'$ sont tangents si et seulement si la restriction de q au plan engendré par H et H' est dégénérée.*

3.6. Solution du problème : alternative de Steiner

Voici la proposition qui va permettre de trivialiser[5] l'alternative de Steiner :

[5]Que l'on nous excuse cette expression un peu triviale, mais l'expression proposée au départ l'était encore plus.

3.6.1. Proposition. *Soit $\mathscr{C}$ et $\mathscr{C}'$ deux cercles-droites qui ne se coupent pas. Alors, il existe $f \in \mathrm{PSL}_2(\mathbb{C})$ tel que $f(\mathscr{C})$ et $f(\mathscr{C}')$ soient deux cercles concentriques.*

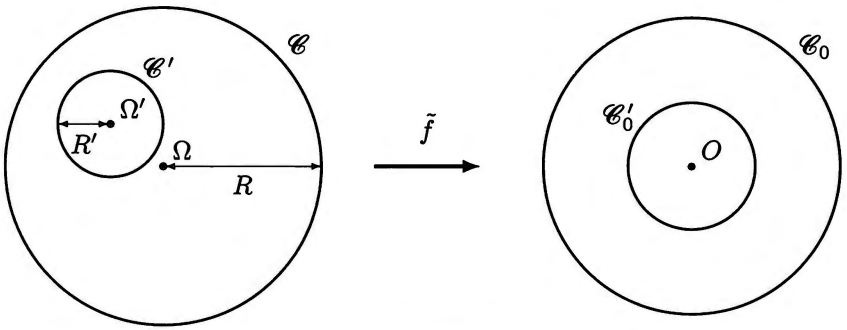


Figure 3.4. *Régularisation de figure pour l'alternative de Steiner*

Démonstration. La stratégie consiste à transposer le problème de double transitivité de $\mathrm{PSL}_2(\mathbb{C})$ sur les cercles, en un problème de double transitivité de $\mathrm{SO}_0(3,1)$ sur les formes hermitiennes. On sait alors que ce type de problème peut se résoudre à l'aide du théorème de Witt.

Pour simplifier, on suppose que $\mathscr{C}$ et $\mathscr{C}'$ sont deux cercles. On reprend les notations du lemme 3.5.6, de sorte que $\mathscr{C} = \mathscr{C}_H$ et $\mathscr{C}' = \mathscr{C}_{H'}$. On y a vu les trois formules : $q(H) = a^2R^2$, $q(H') = a'^2R'^2$, et $2\varphi(H,H') = aa'(R^2 + R'^2 - \Omega\Omega'^2)$.

Comme les cercles $\mathscr{C}$ et $\mathscr{C}'$ ne se coupent pas, on a par le corollaire 3.5.7 : $\varphi(H,H')^2 > q(H)q(H')$.

On peut maintenant mettre en place les hypothèses du théorème de Witt pour envoyer les matrices H, H' respectivement vers des matrices H_0, H'_0 qui correspondront à des cercles concentriques.

On cherche donc deux matrices hermitiennes

$$H_0 = \begin{pmatrix} a_0 & 0 \\ 0 & d_0 \end{pmatrix} \quad \text{et} \quad H'_0 = \begin{pmatrix} a'_0 & 0 \\ 0 & d'_0 \end{pmatrix}$$

telles que $q(H_0) = q(H)$, $q(H'_0) = q(H')$ et $\varphi(H_0, H'_0) = \varphi(H,H')$, d'où le système :

$$\begin{cases} -a_0d_0 = q(H) \\ -a'_0d'_0 = q(H') \\ ad'_0 + a'_0d_0 = -2\varphi(H,H'), \end{cases} \qquad \text{qui donne :} \quad \begin{cases} ad'_0 + a'_0d_0 = -2\varphi(H,H') \\ (a_0d'_0)(a'_0d_0) = q(H)q(H'). \end{cases}$$

Ainsi, $a_0d'_0$ et a'_0d_0 sont les solutions de l'équation de degré 2 suivante :

$$X^2 + 2\varphi(H,H')X + q(H)q(H') = 0.$$

Cette équation a des solutions réelles, car $\varphi(H,H')^2 > q(H)q(H')$. Ainsi, les matrices H_0 et H_0' existent. D'après le théorème de Witt, il existe $\tilde{f} \in \mathrm{O}(q)$ telle que $\tilde{f}(H) = H_0$ et $\tilde{f}(H') = H_0'$.

Il reste à montrer que l'on peut choisir $\tilde{f}$ dans $\mathrm{SO}_0(q)$, quitte à remplacer H_0 et H_0' par leurs opposées, ce qui ne change rien aux cercles correspondants. Comme la restriction de la forme quadratique q au sous-espace engendré par H et H' est de signature $(1,1)$ (puisque son discriminant $q(H)q(H') - \varphi(H,H')^2$ est négatif), le lemme *ad hoc* 3.6.2 s'applique (on pourra se référer au second tome, § VI-3 pour plus de précisions).

L'existence de $\tilde{f} \in \mathrm{SO}_0(q) \simeq \mathrm{SO}_0(3,1)$ qui transforme respectivement H en H_0 et H' en H_0' se traduit par l'existence d'une homographie $f \in \mathrm{PSL}_2(\mathbb{C})$ qui transforme $\mathscr{C}$ en $\mathscr{C}'$ et $\mathscr{C}_0$ en $\mathscr{C}_0'$, deux cercles concentriques de centre l'origine.

L'alternative de Steiner est alors claire par symétrie circulaire. □

3.6.2. Lemme. *Soit q une forme quadratique de signature $(3,1)$ sur un espace vectoriel E de dimension 4. Soit F le plan engendré par deux vecteurs v et v' tels que la signature de q sur F soit $(1,1)$. Soit F_0 le plan engendré par deux vecteurs v_0, v_0' tels que la matrice de $q\,|_F$ dans la base (v,v') est égale à la matrice de $q\,|_{F_0}$ dans la base (v_0,v_0'). Alors, il existe un g dans $\mathrm{SO}_0(q)$ et ϵ dans $\{1,-1\}$ tels que $g(v) = \epsilon v_0$, $g(v') = \epsilon v_0'$.*

Démonstration. On sait par le théorème de Witt qu'il existe $\tilde{g}$ dans $\mathrm{O}(q)$ tel que $\tilde{g}(v) = v_0$, $\tilde{g}(v') = v_0'$.

Passons maintenant à la construction suivante : soit G un orthogonal de F. Alors, on voit que $q\,|_G$ est de signature $(2,0)$ (utiliser par exemple l'unicité de la signature). On peut donc trouver quatre vecteurs v^+, $v^- \in F$, w_1^+, $w_2^+ \in G$ deux à deux orthogonaux tels que $q(v^+) = q(w_1^+) = q(w_2^+) = 1$ et $q(v^-) = -1$. On a donc une isométrie de déterminant -1 qui change w_1^+ en $-w_1^+$ et laisse fixe les trois autres. Cette isométrie laisse donc fixe F. Ainsi, quitte à multiplier à droite $\tilde{g}$ par cette isométrie, on peut supposer que $\tilde{g}$ est dans $\mathrm{SO}(q)$.

Maintenant, soit h tel que $h(v^+) = -v^+$, $h(v^-) = -v^-$, $h(w_1^+) = w_1^+$, $h(w_2^+) = w_2^+$. On sait par la preuve du corollaire VI-A.5 que h est une isométrie de $\mathrm{SO}(q)$ telle que sa classe est non triviale dans $\mathrm{SO}(q)/\mathrm{SO}_0(q)$, lui-même isomorphe au groupe $\{1,-1\}$. De plus, h est une homothétie de rapport -1 sur F. Donc, quitte à multiplier $\tilde{g}$ à droite par h et v, v' par ϵ, on peut supposer que $\tilde{g}$ est dans $\mathrm{SO}_0(q)$. □

3.6.3. Remarque. On peut tenter de généraliser en exercice le lemme précédent. Se référer au § VI-3 du second volume.

3.7. *Post mortem*

La fin de la preuve de l'alternative de Steiner mérite un peu plus d'attention. D'abord, on voit lorsque deux cercles sont concentriques plus de symétries dans la figure que lorsqu'ils ne le sont pas. Plus précisément, les rotations autour du centre commun forment un sous-groupe à un paramètre G d'homographies appartenant au stabilisateur de la paire de cercles, le groupe G étant isomorphe à $\mathrm{SO}(2)$. Pourtant, les deux paires de cercles étant dans la même orbite, leurs stabilisateurs sont conjugués. Cela suggère que l'œil n'est pas aussi accoutumé à voir les homographies que les isométries, ce qui n'est peut-être pas si surprenant.

Une manière alternative de voir que le stabilisateur de cette paire de cercles concentriques (ou plus généralement de deux cercles ne s'intersectant pas) est un sous-groupe isomorphe à $\mathrm{SO}(2)$ est de transporter le problème dans l'espace des matrices hermitiennes. On se ramène à travailler sur un espace quadratique réel $(\mathscr{H}, q)$ muni d'une base orthogonale (H_1, H_2, H_3, H_4) telle que $q(H_1) = q(H_2) = q(H_3) = -q(H_4) = 1$. Les deux cercles ne s'intersectant pas peuvent être assimilés à H_3 et H_4. Le stabilisateur de la paire (H_3, H_4) dans $O(q)$ est alors égal au stabilisateur du plan orthogonal $\langle H_3, H_4\rangle^{\perp} = \langle H_1, H_2\rangle$, et ce stabilisateur n'est rien autre que $O(2)$. Maintenant, le stabilisateur dans $\mathrm{SO}_0(3,1)$, celui qui nous intéresse, est alors clairement isomorphe $\mathrm{SO}(2)$ par la description de $\mathrm{SO}_0(q)$ fournie par le corollaire VI-A.5.

En tout état de cause, formulons la fin de la preuve complètement. La réduction présentée en détail permet de supposer que $\mathscr{C}$ et $\mathscr{C}'$ sont concentriques. On voit que pour tout cercle tangent $\mathscr{C}_0$ à $\mathscr{C}$ et $\mathscr{C}'$, il y a deux cercles tangents à $\mathscr{C}$, $\mathscr{C}'$ et $\mathscr{C}_0$: il s'obtiennent de $\mathscr{C}_0$ par application d'une rotation ρ (dont on peut s'amuser à calculer l'angle en termes des rayons).

Maintenant, dire que la chaîne de cercles se referme, c'est dire qu'il existe $n \geqslant 2$ tel que $\rho^n(\mathscr{C}_0) = \mathscr{C}_0$. Comme G agit sans point fixe sur les cercles tangents à $\mathscr{C}$ et $\mathscr{C}'$, toute la question est de savoir si ρ est d'ordre fini ou pas, et ceci ne dépend pas de $\mathscr{C}_0$.

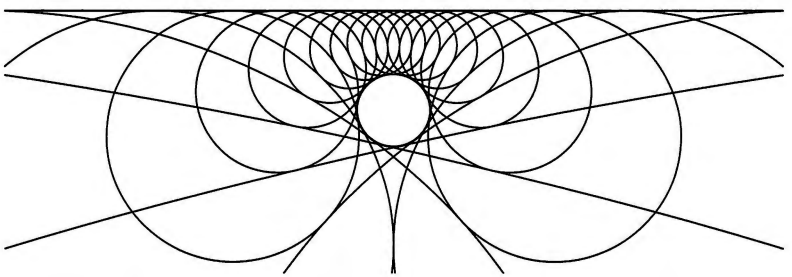
3.7.1. Remarque. L'existence d'un élément d'ordre fini est exactement l'argument au cœur d'une très belle preuve de la plus subtile *alternative de Poncelet*, où le rôle de G est joué par une *courbe elliptique*. On pourra lire l'article original de Griffiths-Harris [17] ou chercher une référence plus accessible quelque part, par exemple l'exposé

`http://www.math.univ-toulouse.fr/~ervedoza/rapport.ps.gz`

Revenons à l'énoncé de l'alternative de Steiner. Certes, on veut bien croire qu'il existe de telles chaînes de cercles, mais est-on sûr de pouvoir toujours

la prolonger ? ou, pour commencer, tracer un premier cercle tangent à $\mathscr{C}$ et $\mathscr{C}'$? La réponse devient limpide après la régularisation de la figure : partant de $\mathscr{C}$ et $\mathscr{C}'$ quelconque, on se ramène à $\mathscr{C}_0$ et $\mathscr{C}'_0$ concentriques par une homographie h ; on construit des chaînes de cercles tangents à $\mathscr{C}_0$ et $\mathscr{C}'_0$; puis on les transforme en chaînes de cercles tangents à $\mathscr{C}$ et $\mathscr{C}'$ par h^{-1}.

Bien plus : dans la présentation initiale du problème, l'hypothèse était qu'un cercle était intérieur à l'autre. Mais l'on n'a jamais utilisé cette hypothèse, sauf pour dire que l'invariant anallagmatique était strictement plus grand que 1, ce qui revient à affirmer que les cercles ne s'intersectent pas. Aussi, on peut étendre la construction de chaînes de cercles à d'autres configurations initiales et partir de deux cercles extérieurs l'un à l'autre, ou bien d'une droite et un cercle : l'alternative de Steiner a toujours lieu.



Références. On ne peut que conseiller [4], chapitres III et V ; pour le birapport, [23], note 0-B.

A. Exercices

Utilisation des groupes de transformations

A.1. Exercice (Centre de gravité d'un triangle)

On considère le plan affine euclidien rapporté à un repère orthonormé. Pour tout couple de vecteurs (u, v), on note $A(u, v)$ l'aire géométrique du parallélogramme engendré par u et v.

1. *Soit g un élément du groupe linéaire du plan vectoriel associé. Montrer l'égalité :*
$$A\big(g(u), g(v)\big) = |\det(g)| A(u, v).$$

2. *On considère un triangle ABC dans le plan et son centre de gravité G. On note A le milieu de $[BC]$, B le milieu de $[AC]$ et C le milieu de $[AB]$. Montrer que les aires géométriques des triangles GAB, GBC, GAC, GBA, GBC et GCA sont égales.*

   On pourra se ramener au cas du triangle équilatéral à l'aide d'un groupe adapté.

A.2. Exercice (Formule de Pick). *Dans le plan $\mathbb{R}^2$ muni de sa structure euclidienne, de sa topologie et sa mesure (notée μ) usuelles, on dit qu'une partie compacte P satisfait à la formule de Pick si*

$$\#(P \cap \mathbb{Z}^2) = \mu(P) + \frac{1}{2}\#(\delta P \cap \mathbb{Z}^2) + 1,$$

où δP désigne la frontière de P.

Le but de l'exercice est de montrer que tout polygone convexe à sommets entiers satisfait à la formule de Pick.

1. *Noter avec admiration que la formule de Pick pour un carré à sommets entiers est donnée par la célébrissime formule $(x+1)^2 = x^2 + 2x + 1$, récemment aperçue dans une publicité pour un réseau de cours à domicile.*
2. *Soit (e, f) une base du réseau $\mathbb{Z}^2$ et soit P le parallélogramme engendré par les vecteurs ae et bf, où a et b sont deux entiers positifs non nuls tels que b divise a.*
   (a) *Montrer que $xe + yf$, $x, y \in \mathbb{R}$ est dans $P \cap \mathbb{Z}^2$ si et seulement si $(x, y) \in \mathbb{Z}^2 \cap [0, a] \times [0, b]$.*
   (b) *Montrer que P satisfait à la formule de Pick.*
   (c) *En déduire que le triangle $T_{ae,bf}$ délimité par les vecteurs ae et bf satisfait à la formule de Pick.*
3. *Déduire par le théorème de la base adaptée via une action de $\mathrm{GL}_2(\mathbb{Z})$ que tout triangle à sommets entiers satisfait à la formule de Pick.*
4. *Conclure en décomposant un polygone convexe à sommets entiers en réunion de triangles à sommets entiers (on fait partir les diagonales d'un sommet du polygone et l'on contemple !).*

A.3. Exercice. *Dans cet exercice, on démontre de façon élémentaire (c'est-à-dire sans le théorème de Witt) que l'on peut envoyer un cercle et un point sur un cercle et son centre par une homographie convenable.*

1. *Soit $\mathscr{S}$ la sphère unité de $\mathbb{R}^3$, $N = (0, 0, 1)$ le « pôle nord », $\mathscr{P}$ un plan qui coupe $\mathscr{S}$ en un cercle $\mathscr{C}$ de rayon strictement positif et le plan tangent à $\mathscr{S}$ en N en la droite Δ. On note I le pôle de Δ par rapport à $\mathscr{C}$ et p la projection centrale de centre N de $\mathscr{P}$ sur le plan équatorial d'équation $z = 0$. Montrer que l'image de $\mathscr{C}$ est un cercle et que l'image de I est le centre de $\mathscr{C}$.*
2. *Soit $\mathscr{C}$ un cercle du plan affine euclidien $\mathbb{R}^2$ et I un point intérieur à $\mathscr{C}$. On veut montrer qu'il existe une homographie $h \in \mathrm{PGL}_3(\mathbb{R})$ qui envoie $\mathscr{C}$ sur un cercle et I sur le centre de $\mathscr{C}'$.*

A.4. Exercice (Théorème du papillon). *L'utilisation de la projection stéréographique pour démontrer ce théorème nous a été indiquée par Bodo Lass.*

1.(a) *Soit $\mathscr{S}$ la sphère unité de $\mathbb{R}^3$, $N = (0,0,1)$ le « pôle nord », $\mathscr{P}$ un plan qui coupe $\mathscr{S}$ en un cercle $\mathscr{C}$ de rayon strictement positif et le plan tangent à $\mathscr{S}$ en N en la droite Δ. On note I le pôle de Δ par rapport à $\mathscr{C}$ (au sens de l'exercice* X-C.15*) et p la projection centrale de centre N de $\mathscr{P}$ sur le plan équatorial d'équation $z = 0$. Montrer que l'image de $\mathscr{C}$ est un cercle et que l'image de I est le centre de $\mathscr{C}$.*

(b) *Soit $\mathscr{C}$ un cercle du plan affine euclidien $\mathbb{R}^2$ et I un point intérieur à $\mathscr{C}$. On veut montrer qu'il existe une homographie $h \in \mathrm{PGL}_3(\mathbb{R})$ qui envoie $\mathscr{C}$ sur un cercle et I sur le centre de $\mathscr{C}'$.*

2. *Soit A et B deux points d'un cercle $\mathscr{C}$, I le milieu du segment $[AB]$. Soit $\mathscr{D}$ et $\mathscr{D}'$ deux droites passant par I. La droite $\mathscr{D}$ (resp. $\mathscr{D}'$) coupe le cercle $\mathscr{C}$ en M et N (resp. M' et N'). On nomme ces points de sorte que M et M' soient dans le même demi-plan de frontière (AB). Soit X l'intersection des droites (AB) et $(M'N)$ (resp. (MN')). Alors, I est le milieu de $[XY]$.*

Incidences, homographies

On note E l'espace vectoriel euclidien $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$, $\mathbb{P} = \mathbb{P}(E)$ et $\mathbb{P}^* = \mathbb{P}(E^*)$.

A.5. Exercice (Incidences)

Ici, $n = 3$. Soit $\mathbb{P}$ un plan projectif $\mathbb{P}(E)$. Nous identifions E à E^* par $\langle\cdot,\cdot\rangle$, donc $\mathbb{P}$ à son dual $\mathbb{P}^*$. Soit D une droite et m un point de $\mathbb{P}$ hors de D. Soit $m^* \subset \mathbb{P}^*$ la droite duale, i.e. *toute droite passant par m. On définit l'*application d'incidence

$$i : m^* \longrightarrow D,$$

en associant, à toute droite passant par m, son point d'intersection avec D. Nous voulons montrer que i est une application projective entre deux droites projectives, i.e. *elle est le projectivisé d'un isomorphisme (linéaire) entre deux plans vectoriel.*

1. *Soit D_m la droite vectorielle de E correspondant au point m. Montrer que m^* s'identifie à $\mathbb{P}(E/D_m)$.*

   Toute droite de $\mathbb{P}$ contenant m s'identifie à un plan vectoriel de E contenant D_m, puis à une droite vectorielle de E/D_m.

2. *Soit P_D le plan vectoriel de E correspondant à la droite D. Montrer que l'ensemble des points de D dans $\mathbb{P}$ s'identifie à l'espace projectif $\mathbb{P}(P_D)$.*

3. *Considérer l'isomorphisme naturel qui à un vecteur de P_D associe sa classe modulo D_m dans E/D_m et conclure.*

A.6. Exercice (Perspectives)

1. *Soit H et H' deux hyperplans de l'espace projectif $\mathbb{P}(E)$, m un point situé ni sur H ni sur H'. Soit x un point de H. Montrer que la droite mx rencontre H' en un point unique, noté $g(x)$. Montrer que g est une application projective.* On s'inspirera à volonté de l'exercice A.5.

   On appelle g la perspective *de centre m de H sur H'.*

2. *Soit D et D' deux droites du plan projectif $\mathbb{P}$ et soit m un point de P situé ni sur D ni sur D'.*

   (a) En identifiant $\mathbb{P}$ avec son dual $\mathbb{P}^*$, montrer que l'application de H sur m^*, associant x à mx, s'identifie à une application d'incidence (exercice A.5*).*

   (b) Montrer que la perspective g de centre m de H sur H' est la composée des applications

$$\begin{array}{ccc} H \longrightarrow m^* & & m^* \longrightarrow H' \\ x \longmapsto mx & \text{et} & d \longmapsto d \cap H'. \end{array}$$

   En déduire que g est aussi composée d'applications d'incidences.

A.7. Exercice (Théorème de Pappus)

Soit $\mathscr{D}$ et $\mathscr{D}'$ deux droites dans un plan projectif et soit A, B et C trois points de $\mathscr{D}$, A', B' et C' trois points de $\mathscr{D}'$. Soit α, β et γ les points d'intersection de $B'C$ et $C'B$, $C'A$ et $A'C$, et $A'B$ et $B'A$ respectivement. Alors, α, β et γ sont alignés.

Soit O le point d'intersection de $\mathscr{D}$ et $\mathscr{D}'$, M celui de BC' et $A'C$, N celui de AC' et $A'B$.

1. *Considérer la composition*

$$f : BC' \longrightarrow \mathscr{D}' \longrightarrow A'B$$

   des perspectives de centre C de BC' sur $\mathscr{D}'$ et de centre A de $\mathscr{D}'$ sur $A'B$. Déterminer $f(B)$, $f(M)$, $f(C')$ et $f(\alpha)$.

2. *Considérer la perspective de centre β*

$$g : BC' \longrightarrow A'B.$$

   Déterminer $g(B)$, $g(M)$ et $g(C')$.

3. *Identifier $g(\alpha)$ et conclure.*

Chapitre XII

Solides platoniciens et sous-groupes finis de $\mathbf{SO(3,\mathbb{R})}$

"And now for something completely different!"

Flying Circus, Monty Python, 1969.

1. Présentation

Voici une première approche des sous-groupes finis de $\mathrm{SO}(3,\mathbb{R})$. Ce chapitre devrait pouvoir se lire indépendamment des autres chapitres de ce livre, tout en restant totalement dans la même thématique.

Il s'agit là d'une leçon simple et très utile pour illustrer l'importance des groupes et des actions de groupes sur des objets concrets de l'espace : les solides platoniciens, c'est-à-dire les polytopes réguliers convexes de notre espace euclidien de dimension trois. L'intérêt de cette leçon réside justement dans le fait que l'on y retrouve dans un cadre concret des notions comme les groupes, les sous-groupes de Sylow, la dualité, les barycentres, les isométries... C'est un exemple phare de leçon transversale entre algèbre et géométrie.

Même si d'un point de vue topologique, les « graphes » de tous les solides platoniciens seront retrouvés via le symbole de Schläfli, nous admettrons leur existence au sens métrique. En fait, leur existence peut être montrée, soit par une succession de données (par exemple les coordonnées des

sommets) pas très illuminante, soit par la classification des sous-groupes finis de $\mathrm{SO}(3,\mathbb{R})$, soit alors à l'aide de la théorie des représentations (plus particulièrement représentations réelles). Le chapitre X du second volume nous montre comment construire le cube (sic !), le tétraèdre, et l'icosaèdre. La dualité (voir la section 4) permet de construire sans peine l'octaèdre (d'ailleurs également obtenu dans le chapitre X du second volume) à partir du cube et enfin le dodécaèdre à partir de l'icosaèdre.

La démarche adoptée dans ce chapitre consiste donc à partir des solides platoniciens et d'en déduire des propriétés des groupes d'isométries associés. Cette démarche est, on le rappelle, indépendante de la théorie des représentations, même si celle-ci est fatalement sous-jacente dans les isomorphismes que nous rencontrerons entre groupes de permutations et groupes d'isométries.

1.1. Définition. Dans $\mathbb{R}^3$, on appelle *solide platonicien* un polyèdre de dimension 3 (d'intérieur non vide) qui est régulier (faces identiques et régulières) et convexe.

1.2. Définition. On appelle *symbole de Schläfli* d'un solide platonicien donné le couple $\{p,q\}$, où p est le nombre d'arêtes de chacune de ses faces et q le degré de chacun de ses sommets (nombre d'arêtes ayant ce sommet pour extrémité).

On admettra (provisoirement) la classification des solides platoniciens, modulo l'action des similitudes. Voici un tableau synthétique de ce que nous allons étudier dans ce chapitre.

Polyèdre X	**Dual**	S	A	F	**Schläfli**	$\mathrm{Is}(X)$	$\mathrm{Is}^+(X)$
Tétraèdre Δ_4	Δ_4	4	6	4	$\{3,3\}$	$\mathfrak{S}_4$	$\mathfrak{A}_4$
Hexaèdre[1] C_6	Δ_8	8	12	6	$\{4,3\}$	$\mathfrak{S}_4 \times \mathbb{Z}/2\mathbb{Z}$	$\mathfrak{S}_4$
Octaèdre Δ_8	C_6	6	12	8	$\{3,4\}$	$\mathfrak{S}_4 \times \mathbb{Z}/2\mathbb{Z}$	$\mathfrak{S}_4$
Dodécaèdre P_{12}	Δ_{20}	20	30	12	$\{5,3\}$	$\mathfrak{A}_5 \times \mathbb{Z}/2\mathbb{Z}$	$\mathfrak{A}_5$
Icosaèdre Δ_{20}	P_{12}	12	30	20	$\{3,5\}$	$\mathfrak{A}_5 \times \mathbb{Z}/2\mathbb{Z}$	$\mathfrak{A}_5$

Figure 1.1. *(S = nombre de sommets; A = nombre d'arêtes; F = nombre de faces)*

[1]Parfois appelé « cube » par l'homme de la rue.

2. Approche topologique

Nous allons voir un résultat étonnant. En dimension trois, la seule contrainte sur les solides platoniciens est une contrainte de nature topologique :

1. de tout sommet part le même nombre d'arêtes ;
2. toute face possède le même nombre d'arêtes ;
3. le solide est homéomorphe à une sphère.

Finalement, une fois assurées ces contraintes topologiques, on verra par un examen au cas par cas peu glorieux, que chaque type de solide régulier topologique se réalise de façon unique en un solide platonicien. En d'autres termes, il existe un solide platonicien pour chaque type ; de plus, si deux solides platoniciens sont homéomorphes (au sens où il existe un homéomorphisme qui envoie sommets sur sommets, arêtes sur arêtes et faces sur faces), alors ils sont semblables. (La situation présente une certaine analogie avec celle des coniques affines réelles : si deux telles coniques sont homéomorphes, alors il existe une bijection affine du plan qui envoie l'une sur l'autre.)

2.1. Définition. On note A le nombre d'arêtes d'un polyèdre, S le nombre de sommets et F le nombre de faces.

On appelle *caractéristique d'Euler-Poincaré* d'un polyèdre la quantité

$$\chi = S - A + F.$$

2.2. Remarque. On peut définir cette quantité pour un grand nombre d'espaces topologiques.

2.3. Théorème (Descartes-Euler). *Tout polyèdre de genre* 0 *a pour caractéristique d'Euler-Poincaré :* $\chi = 2$.

2.4. Remarque. Le genre d'une surface connexe est le nombre maximum de courbes fermées simples ne se coupant pas et pouvant être tracées à l'intérieur de cette surface sans qu'elle perde sa connexité si on la découpait le long de ces courbes. On peut le voir comme le nombre de « trous » sur cette surface. C'est un invariant topologique partiel (deux surfaces de genres différents ne peuvent pas être homéomorphes) et total si l'on ne tient compte que des surfaces orientables.

On se contentera d'annoncer qu'un polyèdre convexe est de genre 0, d'où le corollaire suivant.

2.5. Corollaire. *Pour tout polyèdre convexe, on a :* $S - A + F = 2$.

2.6. Théorème. *Il existe exactement cinq solides platoniciens « topologiques ».*

Démonstration. Les nombres A, S, F, p, q satisfont aux contraintes topologiques suivantes :

$$\begin{cases} 2A = Sq & \text{car à chaque arête correspondent deux sommets}^2; \\ 2A = Fp & \text{car chaque arête est commune à deux faces}; \\ S - A + F = 2; & \\ p \geqslant 3 & \text{car une face est définie à partir d'au moins trois arêtes}; \\ q \geqslant 2 & \text{car aucun sommet n'est l'extrémité d'une arête seulement.} \end{cases}$$

Donc, $\frac{2A}{q} + \frac{2A}{p} = 2 + A$ ce qui donne $\frac{1}{q} + \frac{1}{p} = \frac{1}{A} + \frac{1}{2}$, et comme $p \geqslant 3$, il vient $\frac{1}{q} = \frac{1}{A} + \frac{1}{2} - \frac{1}{p} > \frac{1}{2} - \frac{1}{p} \geqslant \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$ et ainsi, $q < 6$.

Si $q = 2$ alors $\frac{1}{2} + \frac{1}{p} = \frac{1}{2} + \frac{1}{A} \Longrightarrow p = A$ et l'on aurait dans ce cas que deux faces, puisque $2A = Fp$. Un polygone n'est pas un solide platonicien, donc $q \geqslant 3$. Maintenant, p et q jouent des rôles symétriques et donc $p < 6$.

D'où, $3 \leqslant p \leqslant 5$ et $3 \leqslant q \leqslant 5$; on peut vérifier en un nombre fini de cas toutes les possibilités du tableau. □

2.7. Remarque. Pour une construction métrique des solides platoniciens à partir de ces considérations topologiques, il faut ajouter aux contraintes la formule :

$$\sin\frac{\theta}{2} = \frac{\cos\frac{\pi}{q}}{\sin\frac{\pi}{p}},$$

où θ est l'angle entre deux faces adjacentes. On pourra noter que cette formule peut être vue comme la fameuse loi des sinus adaptée aux triangles sphériques.

3. Groupes d'isométries

Le groupe des isométries (affines) d'un objet « mesure » ses symétries (en plus d'avoir une structure de groupe).

[2]Remarque. Plus précisément, ceci provient du fait que les polyèdres sont réguliers (tous les sommets ont même degré) et du lemme des poignées de mains : si l'on note d_i les degrés des sommets, on a $2A = \sum_{i=1}^{S} d_i$.

3.1. Définition. Le groupe $\mathrm{Is}(X)$ des isométries d'un objet $X \subset \mathbb{R}^3$ est le sous-groupe des isométries de l'espace affine euclidien $\mathbb{R}^3$ qui stabilisent X.

Il faut faire attention à ce que l'on dit quand on parle du groupe d'isométries d'un solide platonicien, puisque ce dernier a été défini à similitude près. On va voir que deux objets en similitude ont le même groupe d'isométries à isomorphisme près.

3.2. Proposition. *Soit φ une similitude. Alors :*

$$\mathrm{Is}(X) \simeq \mathrm{Is}\big(\varphi(X)\big).$$

Démonstration. Considérons le morphisme :

$$\begin{array}{ccc} \mathrm{Is}(X) & \longrightarrow & \mathrm{Is}\big(\varphi(X)\big) \\ g & \longmapsto & \varphi g \varphi^{-1}. \end{array}$$

Il s'agit en premier lieu de vérifier qu'il est bien défini. On remarque d'abord que si $g \in \mathrm{Is}(X)$, on a : $\varphi g \varphi^{-1}\big(\varphi(X)\big) = \varphi\big(g(X)\big) = \varphi(X)$. Ensuite, en écrivant φ comme composée d'une homothétie λ et une isométrie ψ, on a que : $\varphi g \varphi^{-1} = (\lambda\psi) g (\lambda\psi^{-1}) = \psi g \psi^{-1}$, qui est bien une isométrie.

Ce morphisme est clairement injectif et surjectif puisque visiblement inversible. □

3.3. Remarque. On s'est implicitement servi du fait que si φ était une similitude, et g une isométrie, alors $\varphi g \varphi^{-1}$ est une isométrie, c'est-à-dire que le groupe des similitudes normalise celui des isométries. En fait, on montre facilement que le groupe des isométries est exactement le normalisateur, dans le groupe affine, du groupe des isométries.

Voici maintenant deux propositions qui vont d'une part ramener l'étude de $\mathrm{Is}(X)$ à celle de $\mathrm{Is}^+(X)$ (le sous-groupe des déplacements de $\mathrm{Is}(X)$), d'autre part ramener l'étude de $\mathrm{Is}^+(X)$ à l'étude de permutations de sommets. On commence pour cela par la définition qui suit.

3.4. Définition. Soit $\mathscr{S}$ un ensemble de points de $\mathbb{R}^n$. On appelle *enveloppe convexe* de $\mathscr{S}$ l'ensemble des barycentres de parties finies de $\mathscr{S}$ dont tous les coefficients sont positifs.

Un point S de l'enveloppe convexe de $\mathscr{S}$ est dit *extrémal* si S n'est pas barycentre à coefficients positifs de $X \setminus \{S\}$.

3.5. Remarque. L'idée dans cette définition est de caractériser par la géométrie affine, qui respecte la notion de barycentre, l'ensemble des sommets d'un polyèdre convexe. Pour l'intuition (on ne s'en servira pas), il n'est pas totalement inutile de rappeler ici le théorème de Krein-Milman : tout convexe compact d'un espace affine de dimension finie est l'enveloppe convexe de ses points extrémaux.

3.6. Proposition. *Soit X une partie finie de $\mathbb{R}^3$.*

(i) *Si O est un centre de symétrie de X et $g \in \mathrm{Is}(X)$, alors : $g(O) = O$. De plus, on a un isomorphisme :* $\mathrm{Is}(X) \simeq \mathrm{Is}^+(X) \times \mathbb{Z}/2\mathbb{Z}$.

(ii) *Si X est stable par une réflexion orthogonale, alors :*

$$\mathrm{Is}(X) \simeq \mathrm{Is}^+(X) \rtimes \mathbb{Z}/2\mathbb{Z}\,.$$

Démonstration. (i) Soit s_0 la symétrie de centre O : l'hypothèse signifie que $s_0 \in \mathrm{Is}(X)$ et, puisque l'on est en dimension 3, s_0 est une isométrie indirecte de X. Puisque toute application affine de l'espace préserve le barycentre, s_0 fixe l'isobarycentre de X. Mais O est le seul point fixe de s_0 donc c'est l'isobarycentre de X. Par suite, g fixe O. Par conséquent : $g \circ s_0 = s_0 \circ g$: en effet, l'application linéaire associée à s_0 est $-\mathrm{Id}$, qui commute avec l'application linéaire associée à g, et $s_0 \circ g$ et $g \circ s_0$ ont un point fixe commun, O.

Mais alors, l'application suivante est un isomorphisme de groupe :

$$\begin{aligned} F : \mathrm{Is}(X) &\longrightarrow \mathrm{Is}^+(X) \times \mathbb{Z}/2\mathbb{Z} \\ g &\longmapsto \begin{cases} (g,1) & \text{si } g \in \mathrm{Is}^+(X) \\ (gs_0, s_0) & \text{sinon.} \end{cases} \end{aligned}$$

(ii) Soit s_0 une réflexion qui préserve X. On dispose de deux sous-groupes : $K = \mathrm{Is}^+(X)$ et $H = \{\mathrm{Id}, s\} \simeq \mathbb{Z}/2\mathbb{Z}$; le premier étant distingué dans $\mathrm{Is}(X)$ et l'on a : $\mathrm{Is}(X) = KH$. En effet, on peut définir une section F de l'application $K \times H \to \mathrm{Is}(X)$, $(k,h) \mapsto k \circ h$ comme au (i) ci-dessus. Par la proposition II-5.2.1, on reconnaît là un produit semi-direct. □

3.7. Proposition. *Soit $X \subset \mathbb{R}^n$. Si $\mathscr{S}$ est un ensemble fini de points et X l'enveloppe convexe de ces points. On suppose que les points de $\mathscr{S}$ sont extrémaux pour X, alors* $\mathrm{Is}(X)$ *stabilise $\mathscr{S}$ et fixe, en particulier, l'isobarycentre de $\mathscr{S}$.*

Démonstration. Le groupe $\mathrm{Is}(X)$ est contenu dans le groupe affine $\mathrm{GA}(\mathbb{R}^3)$, donc conserve les barycentres, c'est-à-dire que si G est le barycentre de la famille (A_i, λ_i), alors $g(G)$ est barycentre de $(g(A_i), \lambda_i)$. Donc, si $g \in \mathrm{Is}(X)$, $S \in \mathscr{S}$, alors $g(S)$ doit être extrémal. En effet, si ce n'était pas le cas, on

aurait $g(S)$ barycentre (non trivial, *i.e.* tous les λ_i ne sont pas concentrés en un seul i) de points de $X\backslash\{g(S)\}$ et comme g^{-1} stabilise le barycentre, on aurait que S est barycentre non trivial de points de $X\backslash\{S\}$. Absurde. □

3.8. Remarque. On admettra sans effort que les points extrémaux d'un solide platonicien sont ses sommets. Ainsi, si g préserve le solide platonicien X, il préserve aussi l'ensemble de ses sommets. On verra en 4.4 qu'il préserve aussi l'ensemble de ses faces et celui de ses arêtes.

3.9. Remarque. Notons que l'on a aussi ramené, à l'aide de la conservation du barycentre, l'étude des isométries affines d'un ensemble fini X possédant un centre de symétrie, à celle des isométries vectorielles. Cette propriété est bien sûr encore valable si X est compact.

3.10. Remarque. Une partie finie X possède un centre de symétrie O si et seulement si la symétrie de centre O préserve X.

Ainsi, puisque le tétraèdre régulier n'a pas de centre de symétrie, son groupe d'isométries ne contient pas de symétrie centrale et $\mathrm{Is}(\Delta_4)$ n'est pas isomorphe au produit direct $\mathrm{Is}^+(\Delta_4) \times \mathbb{Z}/2\mathbb{Z}$. En effet, $\mathrm{Is}(\Delta_4) \simeq \mathfrak{S}_4 \simeq \mathfrak{A}_4 \rtimes \mathbb{Z}/2\mathbb{Z} \simeq \mathrm{Is}^+(\Delta_4) \rtimes \mathbb{Z}/2\mathbb{Z}$, le produit n'est pas direct.

3.11. Exercice. *Montrer que...*
- *si X est un triangle quelconque (pas un quelconque triangle !) du plan, alors* $\mathrm{Is}(X) \simeq \{1\}$.
- *si X est un triangle isocèle, alors* $\mathrm{Is}(X) \simeq \mathbb{Z}/2\mathbb{Z}$.
- *si X est un triangle équilatéral, alors* $\mathrm{Is}(X) \simeq \mathfrak{S}_3$
  Comme $\mathrm{Is}(X)$ stabilise aussi l'ensemble des sommets de X, c'est-à-dire que l'action est fidèle, on a $\mathrm{Is}(X) \subset \mathfrak{S}_3$; et puisque $\mathrm{Is}(X)$ contient tous ses générateurs, on conclut.
- *plus généralement, si X est un polygone régulier à n côtés alors* $\mathrm{Is}(X) \simeq D_n \simeq \mathbb{Z}/n\mathbb{Z} \rtimes \mathbb{Z}/2\mathbb{Z}$. *Attention, on est dans le plan, donc la symétrie centrale n'est pas centrale. On remarquera que* $\mathfrak{S}_3 \simeq D_3$.
  L'oscar de la symétrie revient bien sûr au cercle :
- *si X est un cercle, alors* $\mathrm{Is}^+(X) \simeq \mathbf{S}^1$ *et* $\mathrm{Is}(X) \simeq \mathrm{Is}^+(X) \rtimes \mathbb{Z}/2\mathbb{Z}$. *Ici,* $\mathrm{Is}^+(X)$ *est carrément un groupe de Lie.*

3.12. Proposition. *Les groupes d'isométries d'un tétraèdre régulier Δ_4 sont :*

$$\mathrm{Is}(\Delta_4) \simeq \mathfrak{S}_4 \quad \textit{et} \quad \mathrm{Is}^+(\Delta_4) \simeq \mathfrak{A}_4.$$

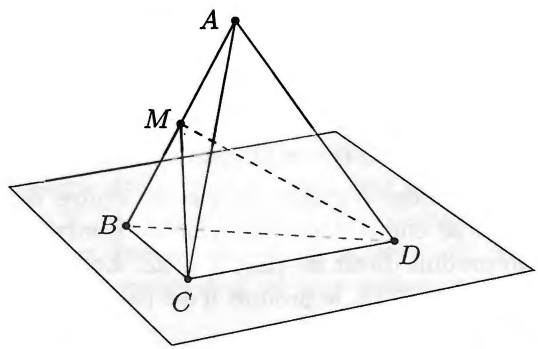
Démonstration. Nous avons admis en remarque 3.8 que les points extrémaux du tétraèdre sont ses sommets. On fait agir $\mathrm{Is}(\Delta_4)$ sur l'ensemble des

sommets du tétraèdre par la proposition précédente : $\mathscr{S} = \{A, B, C, D\}$. On obtient par restriction un morphisme de groupe

$$\begin{aligned} \varphi : \mathrm{Is}(\Delta_4) &\longrightarrow \mathfrak{S}_4 \\ g &\longmapsto g|_{\mathscr{S}}. \end{aligned}$$

L'action est fidèle, c'est-à-dire que φ est injective : si $\varphi(g) = \mathrm{Id}_{\mathscr{S}}$, alors g stabilise $\mathscr{S}$, qui est un repère affine de l'espace, d'où : $g = \mathrm{Id}_{\mathbb{R}^3}$. Ainsi, φ est une injection de $\mathrm{Is}(\Delta_4)$ dans $\mathfrak{S}_4$.

De plus, la réflexion par rapport au plan MCD avec M milieu de AB réalise (a pour image) la transposition $(A\,B)$:



On vérifie de même que toutes les transpositions sont dans l'image par φ de $\mathrm{Is}(\Delta_4)$. Comme les transpositions engendrent $\mathfrak{S}_4$, on en déduit la surjectivité de φ. Finalement, φ est un isomorphisme : $\mathrm{Is}(\Delta_4) \simeq \mathfrak{S}_4$.

Comme $\mathrm{Is}^+(\Delta_4)$ est d'indice 2 dans $\mathrm{Is}(\Delta_4)$, on a aussi : $\mathrm{Is}^+(\Delta_4) \simeq \mathfrak{A}_4$. □

3.13. Remarque. Il s'agit d'une preuve assez simple, mais emblématique de l'utilisation des générateurs pour montrer la surjectivité d'un morphisme, ainsi que de l'utilisation des repères pour montrer l'injectivité.

3.14. Exercice. *Généraliser cette preuve en dimension n : construire un n-simplexe régulier qui généralise le triangle équilatéral en dimension 2 et le tétraèdre en dimension 3, puis montrer que son groupe d'isométries est isomorphe à $\mathfrak{S}_{n+1}$.*

3.15. Proposition. *Les groupes d'isométries du cube sont*

$$\mathrm{Is}^+(C_6) \simeq \mathfrak{S}_4 \quad \textit{et} \quad \mathrm{Is}(C_6) \simeq \mathfrak{S}_4 \times \mathbb{Z}/2\mathbb{Z}.$$

Démonstration. Les grandes diagonales du cube, qui relient deux sommets opposés, sont au nombre de quatre. Ce sont les plus grandes distances

que l'on puisse trouver entre deux points du cube. Donc, une isométrie transforme une grande diagonale en une grande diagonale. (Jeu : trouver un argument permettant de montrer la même chose avec uniquement des arguments affines ? topologiques ?)

On fait donc agir $\mathrm{Is}^+(C_6)$ sur l'ensemble $\mathscr{D} = \{D_1, D_2, D_3, D_4\}$ des grandes diagonales du cube.

Ainsi,

$$\begin{aligned} \varphi : \mathrm{Is}^+(C_6) &\longrightarrow \mathfrak{S}_4 \\ g &\longmapsto g_{|\mathscr{D}} \end{aligned}$$

est un morphisme de groupes, car les distances sont conservées.

On pourra observer la suite de la preuve sur la figure 3.1.

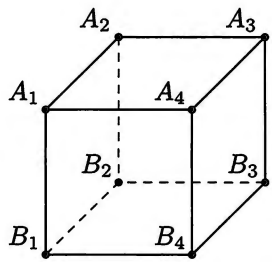


Figure 3.1. *Un cube*

Montrons que l'action que nous venons de définir est fidèle. Soit g tel que $\varphi(g) = \mathrm{Id}_{\mathscr{D}}$ montrons que g est le neutre. Notons D_i la grande diagonale A_iB_i. Comme celle-ci est fixée par g, et que A_i et B_i en sont les points extrémaux, il vient que soit g permute A_i et B_i, soit les laisse tous deux fixes.

Supposons que g laisse fixe A_i (et donc B_i) pour un i donné et laisse fixe globalement les autres diagonales. sans perte de généralité, on peut supposer $i = 1$. Comme $A_1A_2 \neq A_1B_2$ et comme g est une isométrie, il vient que g ne peut envoyer A_2 sur B_2, donc par élimination, A_2 est envoyé sur A_2 et donc B_2 sur B_2. De même, A_4 est envoyé sur lui-même. Or, (A_1, A_2, B_2, A_4) est un repère affine de l'espace. Comme il est fixé par g, l'isométrie g est égale à l'identité.

Si, en revanche, g envoie A_i sur B_i pour i fixé, alors $s_0 g$, où s_0 désigne la symétrie centrale du cube, envoie A_i sur A_i et, d'après ce qui précède, cela implique que $s_0 g$ est l'identité, ce qui est impossible (car g est une isométrie positive).

Conclusion, $\mathrm{Ker}(\varphi) = \{\mathrm{Id}_{\mathbb{R}^3}\}$ et ainsi, comme récompense de la fidélité de l'action, on a : $\mathrm{Is}^+(C_6) \subset \mathfrak{S}_4$.

Comme dans la démonstration précédente, on peut voir que les transpositions sont toutes réalisées (ici grâce à des retournements d'axes reliant les milieux des arêtes joignant les diagonales), et donc $\mathrm{Is}^+(C_6) \simeq \mathfrak{S}_4$. La seconde assertion est claire, car le cube admet un centre de symétrie, d'après la proposition 3.6. □

Dans le cadre des liens robustes entre algèbre et géométrie, on peut noter le phénomène suivant : dans un cube, on peut inscrire deux tétraèdres réguliers T_1 et T_2 dont les sommets sont tous distincts.

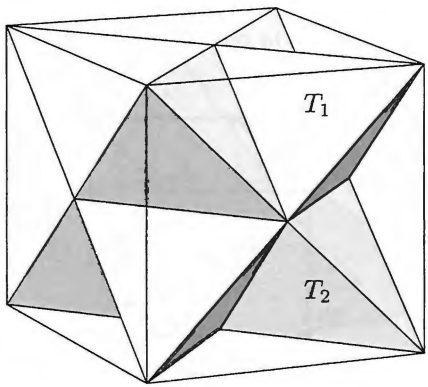


Figure 3.2. *Deux tétraèdres dans un cube*

Le groupe des isométries du cube $\mathrm{Is}(C_6)$ agit sur la paire $\{T_1, T_2\}$ et on peut vérifier que le noyau de l'action est exactement le stabilisateur du tétraèdre :

$$(*) \qquad 1 \longrightarrow \mathrm{Is}(\Delta_4) \hookrightarrow \mathrm{Is}(C_6) \twoheadrightarrow \mathbb{Z}/2\mathbb{Z} \twoheadrightarrow 1$$

D'où, en particulier : $\mathrm{Is}(C_6) \simeq \mathfrak{S}_4 \times \mathbb{Z}/2\mathbb{Z}$, puisque $\mathrm{Is}(\Delta_4) \simeq \mathfrak{S}_4$.

Plus intéressant : on peut aussi restreindre l'action et obtenir la suite exacte :

$$1 \longrightarrow \mathrm{Is}^+(\Delta_4) \longrightarrow \mathrm{Is}^+(C_6) \underset{\leftarrow s -}{\overset{\epsilon}{\longrightarrow}} \mathbb{Z}/2\mathbb{Z} \longrightarrow 1$$

On obtient une section de ϵ à l'aide de la symétrie par rapport à O. Ce qui donne le produit semi-direct $\mathrm{Is}^+(C_6) \simeq \mathrm{Is}^+(\Delta_4) \ltimes \mathbb{Z}/2\mathbb{Z}$, ce qui est équivalent à : $\mathfrak{S}_4 \simeq \mathfrak{A}_4 \ltimes \mathbb{Z}/2\mathbb{Z}$.

Encore plus loin, toujours plus haut, mais avec les mêmes méthodes.

3.16. Proposition. *Les groupes d'isométries du dodécaèdre sont :*

$$\mathrm{Is}(P_{12}) \simeq \mathfrak{A}_5 \times \mathbb{Z}/2\mathbb{Z} \quad \textit{et} \quad \mathrm{Is}^+(P_{12}) \simeq \mathfrak{A}_5.$$

Démonstration. On admet que cinq cubes distincts sont inscrits dans le dodécaèdre.

Pour les construire, un long discours vaut mieux qu'un petit dessin. On part d'une face du dodécaèdre, qui est un pentagone. Alors, à partir d'une diagonale de ce pentagone, on vérifie qu'il y a une unique façon de compléter de sorte qu'elle soit l'arête d'un cube inscrit dans le dodécaèdre. Comme il y a cinq diagonales sur une face, il y a cinq cubes inscrits.

Le groupe $\mathrm{Is}^+(P_{12})$ agit sur l'ensemble $\mathscr{C} = \{C_1, C_2, C_3, C_4, C_5\}$ des cubes inscrits, d'où le morphisme $\mathrm{Is}^+(P_{12}) \longrightarrow \mathfrak{S}_5$.

Soit g tel que $g(C_i) = C_i$ pour tout i. Alors, on voit que g fixe les grandes diagonales du dodécaèdre et comme g n'est pas une symétrie centrale, on obtient comme dans la preuve du groupe du cube que g est l'identité. L'action est donc fidèle et $\mathrm{Is}^+(P_{12}) \subset \mathfrak{S}_5$.

Il est plus simple (même si c'est moins élégant) de montrer la surjectivité par un dénombrement. Combien y a-t-il d'éléments de $\mathrm{Is}^+(P_{12})$? Comme ce sont des rotations, on va compter les axes possibles, puis les angles possibles.

- Axes de sommet à sommet opposé. Il y a $\dfrac{20}{2}$ = dix axes possibles, les angles (non nuls) $\dfrac{2\pi}{3}, \dfrac{4\pi}{3}$.
- Axes de milieu d'arête à milieu d'arête opposée. Il y a $\dfrac{30}{2}$ = quinze axes possibles, les angles (non nuls) π.
- Axes de centre de face à centre de face opposé. Il y a $\dfrac{12}{2}$ = six axes possible, les angles (non nuls) $\dfrac{2\pi}{5}, \dfrac{4\pi}{5}, \dfrac{6\pi}{5}, \dfrac{8\pi}{5}$.
- Et l'identité, bien sûr !

En tout, cela nous fait $10 \times 2 + 15 \times 1 + 6 \times 4 + 1 =$ soixante éléments. Le compte est bon, et $\mathrm{Is}^+(P_{12}) = \mathfrak{A}_5$.

Enfin, le dodécaèdre ayant un centre de symétrie, on conclut à l'aide de la proposition 3.6 que l'on a : $\mathrm{Is}(P_{12}) \simeq \mathfrak{A}_5 \times \mathbb{Z}/2\mathbb{Z}$. □

3.17. Remarque. Des preuves qui précèdent, on peut dégager un procédé plus général pour montrer qu'un groupe donné s'envoie sur $\mathfrak{S}_n$. La bonne question à se poser devant ce genre de problème est : « Le groupe en question agit sur n quoi ? »

On a vu que $\mathrm{Is}(T_4)$ agit sur les quatre sommets, que $\mathrm{Is}^+(C_6)$ sur les quatre grandes diagonales du cube et enfin, plus délicat, que $\mathrm{Is}^+(P_{12})$ agit sur les cinq cubes inscrits dans P_{12}.

L'exercice suivant est caractéristique.

3.18. Exercice. *Montrer qu'il y a un morphisme surjectif de $\mathfrak{S}_4$ vers $\mathfrak{S}_3$.*

Réponse. *Après avoir noté le caractère exceptionnel d'un tel morphisme — vu que, pour $n \geqslant 5$, le groupe $\mathfrak{S}_n$ n'admet pour seules images homomorphes que le groupe trivial, $\mathbb{Z}/2\mathbb{Z}$ et lui-même — on va se poser la question-clé : le groupe $\mathfrak{S}_4$ (c'est-à-dire le groupe du tétraèdre) agit sur trois quoi ? Le tétraèdre ayant quatre faces, six arêtes et quatre sommets, ce sont les arêtes qui nous intéressent. D'autant plus que deux arêtes se touchent sauf si elles sont opposées, propriété préservée par une symétrie du tétraèdre. Donc, le groupe $\mathfrak{S}_4$ agit sur les trois paires d'arêtes opposées du tétraèdre. Pour montrer la surjectivité, il suffit de voir que les transpositions sont atteintes (en fait par symétrie, une seule suffit). On laisse cette partie au lecteur.*

Question. En réalisant $\mathfrak{S}_4$ comme groupe d'isométries positives du cube, sur quel ensemble à trois éléments agit-il naturellement ?

4. La *toy* dualité

Nous allons définir une notion de dualité adaptée à la situation des polyèdres réguliers convexes de $\mathbb{R}^3$. C'est une dualité de pacotille, qui va juste nous permettre de comprendre ce que l'on a besoin de comprendre dans l'immédiat, mais avec laquelle on ne va pas bien loin. On l'appellera *toy* dualité[3] pour la distinguer de la vraie dualité chez les ensembles compacts convexes de $\mathbb{R}^n$.

Pour le lecteur assidu, voire le candidat à l'agrég., nous parlerons de cette dualité plus profonde en appendice. Nous y généraliserons les propriétés étudiées et nous verrons pourquoi la *toy* dualité en est un cas particulier.

4.1. Définition. Soit X solide platonicien de $\mathbb{R}^3$ de centre O. Notons C_i, pour $1 \leqslant i \leqslant f$, les centres de ses faces. Le dual X^* de X est l'enveloppe convexe des C_i.

Si X est un solide platonicien, alors $X^{**} = X$ comme on peut le voir sur le tableau des solides platoniciens.

[3] Cette terminologie n'est valable que dans le cadre de l'ouvrage et réellement assumée que par le plus funky des deux auteurs.

4.2. Proposition. *Soit X un solide platonicien. Alors, on a l'isomorphisme* $\mathrm{Is}(X^*) \simeq \mathrm{Is}(X)$.

Démonstration.
$\boxed{\Longleftarrow}$ Soit X un solide platonicien et $g \in \mathrm{Is}(X)$. Alors, g fixe l'ensemble de sommets de X^* puisqu'il préserve le barycentre. Donc, $\mathrm{Is}(X) \subset \mathrm{Is}(X^*)$.
$\boxed{\Longrightarrow}$ Par dualité, $\mathrm{Is}(X^*) \subset \mathrm{Is}(X^{**}) \simeq \mathrm{Is}(X)$.

□

4.3. Remarque. On peut maintenant déduire par dualité le groupe de l'octaèdre (dual du cube) et celui de l'icosaèdre (dual du dodécaèdre).

On peut se dire que le résultat qui suit va de soi, et il n'est pas défendu de l'admettre en utilisant le sens commun. On en donne tout de même une preuve qui utilise la proposition précédente.

4.4. Corollaire. *Soit X un solide platonicien et $g \in \mathrm{Is}(X)$, alors g préserve l'ensemble des faces et l'ensemble des arêtes de X.*

Démonstration. On « rappelle » que si F est une face de X, alors elle est entièrement déterminée par une équation $P \in F \Leftrightarrow P \in X$ et $l(P) = k$, où l est une forme linéaire et k un scalaire. De plus, on peut choisir l (quitte à changer en $-l$) telle que tout point P de X vérifie $l(P) \geqslant k$, avec égalité si et seulement si P est sur F.

Si F est l'enveloppe convexe des sommets de X, notés $S_1, \ldots, S_r$, et si T est l'isobarycentre de ces sommets, alors, d'après la proposition précédente, le point T est envoyé par g sur un sommet de X^* et donc sur l'isobarycentre d'une face F' de X qui a pour équation, disons, $l'(P) = k'$, avec de plus $l'(P) \geqslant k'$ pour tout P de X. Or, $T' = g(T)$ est aussi l'isobarycentre de $g(S_1), \ldots, g(S_r)$. Par l'inégalité précédente. Comme on sait que la moyenne de nombres supérieurs à k' ne vaut k' que si tous ces nombres sont égaux à k', on obtient que $l'\big(g(S_i)\big) = k'$ pour tout i. Donc, les $g(S_i)$ sont sur la face F'. Et, comme les $g(S_i)$ sont des sommets (distincts), ils engendrent clairement la face F' puisqu'il y en a le bon nombre, par régularité de X.

Pour l'ensemble des arêtes, il suffit de voir qu'une arête est l'intersection dans X de deux faces adjacentes. □

4.5. Définition. Soit X un solide convexe de $\mathbb{R}^3$. On peut aussi définir son dual X^* par : $X^* = \{v \in \mathbb{R}^3 \ : \ \langle u, v\rangle \leqslant 1, \ \forall u \in X\}$.

C'est une dualité qui tient la route. Il est conseillé de regarder en annexe les propriétés de cette dualité, qui a pour cas particulier la *toy* dualité dans le cadre des polyèdres réguliers convexes.

A. Annexe. Dualité des ensembles compacts convexes de $\mathbb{R}^n$

Je suis votre miroir, la belle, réfléchissez pour moi, je réfléchirai pour vous.

La Belle et la Bête, Jean Cocteau, 1946.

On va commencer par un mini-Hahn-Banach (dans le sens que l'on est en dimension finie). Dans la suite, E désignera l'espace $\mathbb{R}^n$ muni de sa structure euclidienne canonique.

A.1. Proposition. *Soit C un convexe compact de l'espace euclidien E et soit a dans E, hors de C. Alors, la fonction qui à tout u de C associe sa distance $||u-a||$ possède un minimum $d>0$ atteint en u_0 dans C.*

Soit $\mathscr{H}$ l'hyperplan affine défini par

$$\mathscr{H} := \{u \in E,\ (u-a)\cdot(u_0-a) = \frac{d^2}{2}\}.$$

Alors, a est dans un des demi-espaces ouverts définis par $\mathscr{H}$ et C est dans l'autre.

Démonstration. L'existence d'une distance minimum vient de la compacité de C et de l'existence du minimum (atteint !) d'une fonction continue sur un compact[4]. Cette distance est strictement positive puisque, sinon, on aurait $a \in C$, ce qui est contraire à l'hypothèse.

Montrons la dernière assertion. On note

$$\mathscr{H}^+ := \{u \in E,\ (u-a)\cdot(u_0-a) > \frac{d^2}{2}\},$$
$$\mathscr{H}^- := \{u \in E,\ (u-a)\cdot(u_0-a) < \frac{d^2}{2}\}$$

les deux demi-espaces ouverts définis par $\mathscr{H}$. Clairement, a est dans $\mathscr{H}^-$. Montrons par l'absurde que C est inclus dans $\mathscr{H}^+$. Pour cela, supposons qu'un vecteur u_1 de C soit dans $\mathscr{H}^-$ et posons $k := (u_1-a)\cdot(u_0-a) < \frac{d^2}{2}$. Alors, la fonction $h(u) := (u-a)\cdot(u_0-u_1)$ est continue sur le segment $[u_1,u_0]$ qui est inclus dans C. Comme

$$h(u_1) = (u_1-a)\cdot(u_0-u_1) = (u_1-a)\cdot\big((a-u_1)+(u_0-a)\big)$$
$$= -||u_1-a||^2 + k < -\frac{d^2}{2},$$

et comme l'on a :

$$h(u_0) = (u_0-a)\cdot(u_0-u_1) = (u_0-a)\cdot\big((a-u_1)+(u_0-a)\big) = -k+d^2 > \frac{d^2}{2}.$$

[4]D'ailleurs, la convexité implique l'unicité de u_0. (Laissé en exercice.)

Par le théorème des valeurs intermédiaires sur h, on a l'existence d'un v dans $[u_1, u_0]$ (donc dans C) tel que $h(v) = 0$. Il résulte de $(v-a)\cdot(u_0-u_1)=0$ que v est le projeté orthogonal de A sur M_1M_0. On a donc $(u_0-a)^2 = (v-a)^2 + (u_0-v)^2$, donc $||v-a|| < ||u_0-a||$, ce qui est impossible. □

Voici maintenant une notion plus sérieuse, et surtout plus maniable, de dualité.

A.2. Définition. Soit C un convexe de E. On définit son dual

$$C^* := \{u \in E,\ u\cdot v \leqslant 1,\ \forall v \in C\}.$$

On peut bien sûr transposer ce dual dans E^* via la forme euclidienne et ainsi voir C^* dans le dual de E.

A.3. Proposition. *Soit C un convexe de E.*

(i) *La partie C^* est convexe et fermée.*

(ii) *Si C contient 0 dans son intérieur, alors C^* est compact.*

(iii) *Si C est compact avec 0 en son intérieur, alors $(C^*)^* = C$.*

Démonstration.

1. la partie C^* est une intersection de demi-espaces fermés, donc est un convexe fermé.
2. Posons ϵ tel que la boule ouverte $B(O,\epsilon)$ de centre O et de rayon ϵ est incluse dans C. Montrons alors que $C^* \subset B(O, \frac{1}{\epsilon})$.
   En effet, soit $u \in C^*$ non nul, alors en particulier, $u\cdot v \leqslant 1$, $\forall v \in B(O, \frac{\epsilon}{2})$. Posons $v := \frac{\epsilon}{2}\frac{u}{||u||}$, alors $u\cdot\frac{\epsilon}{2}\frac{u}{||u||} \leqslant 1$. Donc, en prenant la norme, on obtient $||u|| \leqslant 2$. La partie C^* est donc fermée bornée, donc compacte.
3. Il est clair par construction que $C \subset (C^*)^*$. Montrons l'inclusion inverse par la contraposée.
   Soit $a \notin C$, montrons que $a \notin (C^*)^*$. Soit $\mathscr{H}$ un hyperplan séparant a et C comme dans la proposition précédente. Alors, $0 \notin \mathscr{H}$. L'équation de $\mathscr{H}$ peut donc s'écrire sous la forme $y_0\cdot x = b$ pour un b non nul et donc on peut ainsi trouver x_0 tel que $\mathscr{H} = \{x,\ x_0\cdot x = 1\}$. Tout x de C vérifie $x_0\cdot x < 1$, et, de plus, on a $x_0\cdot a > 1$, par construction. La première inégalité prouve que $x_0 \in C^*$ et la seconde que $a \notin (C^*)^*$, ce qui était l'assertion voulue. □

Comme on l'a vu dans la section sur la *toy* dualité, il est intéressant de constater que les groupes d'isométries de deux compacts convexes duaux sont isomorphes, et même égaux !

A.4. Proposition. *Soit C un compact convexe de $\mathbb{R}^n$ avec 0 en son intérieur, alors*

$$\mathrm{Is}(C) = \mathrm{Is}(C^*).$$

Démonstration. Soit u dans $\mathrm{GL}_n(\mathbb{R})$, alors $u(C)$ est un compact convexe et l'on montre facilement que $u(C)^* = (u^*)^{-1}(C^*)$. Donc, si $u \in \mathrm{O}(n)$, il vient $u(C)^* = u(C^*)$. Cela implique que $\mathrm{Is}(C) \subset \mathrm{Is}(C^*)$. Comme C est compact avec 0 en son intérieur, il en est de même pour $u(C)$. On peut dualiser la relation $u(C)^* = u(C^*)$ pour obtenir $u(C^*)^* = \big(u(C)^*\big)^* = u(C)$ et ainsi obtenir l'inclusion inverse $\mathrm{Is}(C^*) \subset \mathrm{Is}(C)$. □

Il reste enfin à montrer que cette dualité est bien une généralisation de celle que nous avons utilisée sur les polyèdres réguliers convexes de $\mathbb{R}^3$.

A.5. Proposition. *Dans l'espace affine de dimension 3 muni d'un repère orthonormé d'origine O (et ainsi identifié à l'espace vectoriel $\mathbb{R}^3$), soit P un polyèdre convexe compact régulier de centre 0. Notons C_i, pour $1 \leqslant i \leqslant f$, les centres de ses faces et soit P' l'enveloppe convexe des C_i. Supposons $OC_i = 1$ pour tout i. Alors, on a $P^* = P'$.*

Démonstration. Comme $OC_i^2 = 1$, les faces du polyèdre P ont pour équation $\overrightarrow{OM} \cdot \overrightarrow{OC_i} = 1$ pour tout i, donc $P = \{M,\ \overrightarrow{OM} \cdot \overrightarrow{OC_i} \leqslant 1,\ 1 \leqslant i \leqslant f\}$.

Maintenant, le dual de P' est donné par l'inéquation $\overrightarrow{OM} \cdot \overrightarrow{ON} \leqslant 1$ pour tout N dans P', et donc $\overrightarrow{OM} \cdot (\sum_i \lambda_i \overrightarrow{OC_i}) \leqslant 1$ pour tout f-uplet de λ_i positifs. Cela revient à dire $\overrightarrow{OM} \cdot \overrightarrow{OC_i} \leqslant 1$ pour tout i.

On obtient donc par ce qui précède $P = (P')^*$ et comme P est compact avec O en son intérieur, $P' = P^*$, comme demandé. □

Références. Pour ce paragraphe, on peut se référer au livre [13], paragraphes 5.1 et 5.2, ou aux exercices partiellement corrigés du fameux livre [4], exercices IV.27, IV.28, IV.29, IV.30.

B. Sous-groupes de Sylow d'un groupe d'isométries

Les groupes d'isométries des solides platoniciens ont l'avantage d'avoir des p-Sylow « visibles à l'oeil nu ». Voici quelques exemples témoins de l'interaction omniprésente entre groupes et géométrie.

B.1. Exemple. Combien $\mathfrak{A}_4$ possède-t-il de 3-Sylow ?

Bien sûr, nous avons la méthode arithmétique qui consiste à regarder les valeurs possibles 1 et 4, et à en éliminer une par des considérations plus ou moins orthodoxes. Mais ici, nous pouvons utiliser une méthode géométrique, puisque $\mathfrak{A}_4$ se réalise comme groupe d'isométries directes du tétraèdre. Comme ses 3-Sylow ont pour ordre 3, ce sont des rotations d'ordre 3 et la seule possibilité est la rotation autour d'un axe passant par le milieu des faces. Il y a donc quatre 3-Sylow correspondant aux quatre faces d'un tétraèdre.

B.2. Exemple. Combien $\mathfrak{S}_4$ possède-t-il de 3-Sylow ?

Même méthode : cette fois, le groupe $\mathfrak{S}_4$ se réalise comme groupe d'isométrie directe de deux manières : celui du cube et celui de l'octaèdre. Il semble que visuellement, c'est en tant que groupe de l'octaèdre que l'on verra le mieux les 3-Sylow, car ils sont en bijection avec ses paires de faces opposées (triangulaires).

B.3. Exemple. Combien $\mathfrak{A}_5$ possède-t-il de 3-Sylow ?

Réponse. Même méthode ! Il y en a 10, correspondant aux paires de faces opposées de l'icosaèdre.

B.4. Exemple. Combien $\mathfrak{A}_5$ possède-t-il de 5-Sylow ?

Réponse. Même méthode. Six, correspondant aux paires de faces opposées du dodécaèdre.

B.5. Exercice. *Combien $\mathfrak{S}_4$ possède-t-il de 2-Sylow ?*

Réponse. Trois ! Pourquoi ? (géométriquement bien sûr avec les paires de faces opposées).

Montrer qu'ils sont engendrés par un 4-cycle et une transposition et sont donc isomorphes au groupe diédral D_4 (d'ordre 8).

C. Exercices

C.1. Exercice

1. *Pourquoi est-il évident* a priori[5] *que tous les carrés ont des groupes d'isométries isomorphes ?*

   On vérifiera que le groupe des similitudes normalise le groupe des isométries du plan. Quelle est l'orbite d'un carré sous le groupe des similitudes ?

[5] C'est-à-dire, avant de le déterminer.

2. *Pourquoi peut-on affirmer* a priori *que le groupe d'isométries de n'importe quel parallélogramme (ou rectangle ou losange) est un sous-groupe du groupe des isométries d'un carré ?*

   Un parallélogramme est un carré qui s'ignore...

3. *Pourquoi n'est-il pas clair* a priori *que les groupes d'isométries de tous les rectangles sont isomorphes ? Quel est le groupe des isométries d'un rectangle ? Montrer* a priori *que le groupe d'isométries de n'importe quel losange est le groupe d'isométries d'un rectangle convenable.*

C.2. Exercice. *Montrer en utilisant les isométries positives du cube, ou le groupe des isométries du tétraèdre régulier, qu'il existe une surjection de $\mathfrak{S}_4$ sur $\mathfrak{S}_3$.*

On fait agir le groupe du cube sur les paires de faces opposées ou sur les trois axes du cube.

C.3. Exercice

1. *Remarquer en lisant le tableau de solides platoniciens en page 358 que la formule $|\mathrm{Is}^+(X)| = Sq$ est valable pour tout solide platonicien X. Expliquez cette formule en faisant agir transitivement le groupe des déplacements de X sur l'ensemble de ses sommets.*
2. *Par dualité montrer que $|\mathrm{Is}^+(X)| = Fp$.*
3. *Comment expliquer la formule $|\mathrm{Is}^+(X)| = 2A$?*

C.4. Exercice (Une définition forte de régularité). *Soit P un polyèdre régulier. On note $\mathscr{S}$ (resp. $\mathscr{A}$, $\mathscr{F}$) l'ensemble de ses sommets (resp. arêtes, faces). Il y a une relation d'incidence entre les sommets et les arêtes (l'appartenance) et entre les arêtes et les faces (l'inclusion).*

Inspiré par la définition d'un drapeau dans un espace vectoriel et en anticipation d'une remarque du § IV-1 du second volume, on définit un drapeau *de P comme une suite (s, a, f), où s est un sommet appartenant à une arête a, elle-même contenue dans une face f. Par la proposition 4.4, une isométrie d'un polyèdre régulier préserve ses sommets, ses arêtes et ses faces : par suite, il agit sur l'ensemble des drapeaux.*

Par ailleurs, on appelle automorphisme combinatoire *de P une permutation de chacun des ensembles $\mathscr{S}$, $\mathscr{A}$ et $\mathscr{F}$ qui préserve l'incidence, cela étend la notion du second tome, chapitre V, lorsque l'on a trois catégories d'objets.*

1. *Démontrer que le groupe des isométries de P agit de façon transitive sur les drapeaux. Inversement, pourquoi est-ce que toute isométrie qui préserve les drapeaux préserve le polyèdre ?*

2. *Déterminer le groupe des automorphismes combinatoires de P.*
3. *Montrer que tout automorphisme combinatoire de P provient d'une isométrie qui préserve P.*

[Motivation : *les* immeubles de Bruhat-Tits *sont des objets géométriques bien plus généraux que les polyèdres. La condition de transitivité du groupe d'automorphismes sur les drapeaux y a un sens et étend, par conséquent, la notion de polyèdre régulier aux immeubles.*]

C.5. Exercice. *L'entraîneur de l'équipe de France de football a décidé de mettre à l'honneur son équipe en créant un ballon de foot pavé uniquement d'hexagones et de représenter ainsi géométriquement un symbole fort de son pays. On demande bien sûr que le ballon soit sphérique, que chaque sommet du pavage participe à au moins trois hexagones, mais l'on ne demande aucune propriété de régularité (les hexagones peuvent être quelconques et le nombre d'hexagones par sommet n'est pas supposé constant). Soit S, A, F le nombre de sommets, d'arêtes, de faces du pavage.*

1. *Montrer l'inégalité $3S \leqslant 2A$.*
2. *Trouver une relation entre A et F qui exprime que chaque hexagone possède six côtés.*
3. *Trouver une relation entre S, A et F qui tient compte du fait que la caractéristique d'Euler-Poincaré est un invariant topologique.*
4. *En conclure à l'impossibilité d'une telle entreprise.*

C.6. Exercice (Nombre de coloriages d'un solide platonicien à isométrie près). *Soit X un solide platonicien. On veut colorier ses faces avec n couleurs fixées (les couleurs ne sont pas forcément toutes représentées et elles peuvent être représentées plusieurs fois). On dit que deux solides du même type ont même modèle de coloriage si l'on peut passer de l'un à l'autre par une rotation de l'espace. On veut compter le nombre de modèles différents de coloriage que l'on peut créer.*

1. *À titre d'exemple, montrer à la main que l'on peut colorier un tétraèdre avec deux couleurs de cinq modèles différents.*
2. *Soit Y l'ensemble de couleurs de cardinal n.*
   (a) *Soit Φ l'ensemble des faces de X. Montrer que colorier X avec les couleurs de Y revient à se donner une fonction de Φ dans Y. On identifie alors l'ensemble des coloriages avec l'ensemble $\mathscr{F}(\Phi, Y)$ des fonctions de Φ dans Y.*
   (b) *Observer que deux coloriages f et f' de $\mathscr{F}(\Phi, Y)$ ont même modèle si et seulement si il existe g de $\mathrm{Is}^+(X)$ tel que $g.f = f'$, où l'action de $\mathrm{Is}^+(X)$ sur $\mathscr{F}(\Phi, Y)$ est celle de la remarque* A.1.13.

(c) En déduire que le nombre de modèles cherché est égal au cardinal de l'ensemble $\mathscr{F}(\Phi,Y)/\mathrm{Is}^+(X)$ des orbites de l'action précédente.

3. En utilisant la formule de Burnside et l'exercice I-A.1.14, *montrer que*

$$|\mathscr{F}(\Phi,Y)/\mathrm{Is}^+(X)| = \frac{1}{|\mathrm{Is}^+(X)|}\sum_{g\in \mathrm{Is}^+(X)} n^{|\Phi/\langle g\rangle|}.$$

4. Retrouver les cinq modèles de coloriage du tétraèdre à deux couleurs fixées.

Le calcul donne $\frac{1}{12}(2^4+8\times 2^2+3\times 2^2)$, où, dans la somme, on retrouve l'élément neutre, les huit 3-cycles et les trois produits de 2-cycles du groupe $\mathscr{A}_4$.

5. Combien de modèles pour le coloriage d'un cube avec trois couleurs ?

Le calcul donne $\frac{1}{24}(3^6+8\times 3^2+6\times 3^3+3\times 3^4+6\times 3^3)=57$, où, dans la somme, on retrouve l'élément neutre, les huit 3-cycles, les six transpositions, les trois produits de deux cycles et enfin les six quatre-cycles de $\mathscr{S}_4$.

Tony Stark : All right, hey ! All right, good job guys ! Let's just not come in tomorrow ; let's just take a day.

The Avengers, Joss Whedon, 2012.

Bibliographie

[1] J. F. ADAMS : Vector fields on spheres. *Ann. of Math. (2)*, 75:603–632, 1962.

[2] Aurélien ALVAREZ, Étienne GHYS et Jos LEYS : *Dimensions. Une promenade mathématique*, 2008.
Film d'animation et site web : `http://dimensions-math.org/`

[3] Yves ANDRÉ : Idées galoisiennes. *In Histoire de mathématiques*, pages 1–16. Ed. Éc. Polytech., Palaiseau, 2012.

[4] Michèle AUDIN : *Géométrie.* EDP-Sciences, Grenoble, deuxième édition, 2006.

[5] Michèle AUDIN : *Analyse complexe*, 2011. Notes d'un cours donné à l'Université de Strasbourg,
`http://www-irma.u-strasbg.fr/~maudin/analysecomp.pdf`

[6] André AVEZ : *Calcul différentiel.* Maîtrise de mathématiques pures. Masson, Paris, 1999.

[7] Rajendra BHATIA : Linear algebra to quantum cohomology : the story of Alfred Horn's inequalities. *Amer. Math. Monthly*, 108(4):289–318, 2001.

[8] Laurent BONAVERO et Jean-Pierre DEMAILLY : *Fonctions holomorphes et surfaces de Riemann*, 2009. Notes (très augmentées) d'un cours donné à l'École normale supérieure de Lyon en 1995-1997 et 2003-2005,
`http://www-fourier.ujf-grenoble.fr/~demailly/manuscripts/`

[9] Jean-Luc CHABERT : Gauss et la méthode des moindres carrés. *Rev. Histoire Sci.*, 42(1-2):5–26, 1989. La mathématisation 1780–1830.

[10] Henri COHEN : *A Course in Computational Algebraic Number Theory.* Graduate Texts in Mathematics, Vol. 138. Springer Verlag, deuxième édition, 1993.

[11] Pierre COLMEZ : *Éléments d'analyse et d'algèbre (et de théorie des nombres).* École Polytechnique (éditions), 2009.

[12] Michel DEMAZURE : *Catastrophes et bifurcations.* Ellipses, 1998.

[13] Jean FRESNEL : *Méthodes modernes en géométrie.* Hermann, Paris, troisième, revue et augmentée édition, 2010.

[14] William FULTON : *Young tableaux*, volume 35 de *London Mathematical Society Student Texts.* Cambridge University Press, Cambridge, 1997. With applications to representation theory and geometry.

[15] Rémi GOBLOT : *Algèbre linéaire.* Mathématiques à l'université. Ellipses, deuxième édition, 2005.

[16] Claude GODBILLON : *Élments de topologie algébrique.* Méthodes. Hermann, 1998.

[17] Phillip GRIFFITHS et Joseph HARRIS : On Cayley's explicit solution to Poncelet's porism. *Enseign. Math. (2)*, 24(1-2):31–40, 1978.

[18] Jean-Pierre KAHANE, éditeur. *L'enseignement des sciences mathématiques. Commission de réflexion sur l'enseignement des mathématiques.* Odile Jacob, Paris, 2002.

[19] Benoît KLOECKNER : *Petit aperçu de géométrie projective.* Calvage et Mounet, 2012.

[20] Franz LEMMERMEYER : *Reciprocity laws.* Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2000. From Euler to Eisenstein.

[21] Ludovic MARQUIS : Géométrie de Hilbert. *Images des Math.*, 2011.

[22] Jean-Yves MÉRINDOL : *Nombres et algèbre.* Collection Grenoble sciences. EDP-Sciences, Grenoble, 2006.

[23] Rached MNEIMNÉ : *Éléments de géométrie.* Nouvelle Bibliothèque Mathématique, 2. Cassini, Paris, 1997. Actions de groupes.

[24] Rached MNEIMNÉ : *Réduction des endomorphismes – Tableaux de Young, cône nilpotent, représentations des algèbres de Lie semi-simples.* Tableau noir. Calvage et Mounet, 2006.

[25] Rached MNEIMNÉ et Frédéric TESTARD : *Introduction à la théorie des groupes de Lie classiques.* Collection Méthodes. Hermann, Paris, 1986.

[26] Sophie MORIER-GENOUD et Valentin OVSIENKO : Well, papa, can you multiply triplets ? *Math. Intelligencer*, 31(4):1–2, 2009.

[27] Hiraku NAKAJIMA : *Lectures on Hilbert schemes of points on surfaces*, volume 18 de *University Lecture Series.* American Mathematical Society, Providence, RI, 1999.

[28] Ivan NOURDIN : *Agrégation de mathématiques.* Collection Sciences Sup. Dunod, deuxième édition, 2006. Épreuve orale – 68 thèmes pour se préparer efficacement.

[29] Daniel PERRIN : *Cours d'algèbre.* Collection CAPES/agrégation. Éditions Ellipses, Paris, 1997.

[30] Felice RONGA : Le calcul de Schubert selon Schubert. Prépublication : `http://arXiv.org/abs/math/0409281v1`, 2004.

[31] Pierre SAMUEL : *Géométrie projective.* Mathématiques. Presses Universitaires de France, Paris, 1986.

[32] Michelle SCHATZMAN : *Analyse numérique – Une approche mathématique.* Dunod, deuxième, revue et augmentée édition, 2001.

[33] Clément de SEGUINS-PAZZIS : *Invitation aux formes quadratiques.* Mathématiques en devenir. Calvage et Mounet, Paris, 2011.

[34] Jean-Pierre SERRE : *Cours d'arithmétique.* Presses Universitaires de France, Paris, 1977. Deuxième édition revue et corrigée, Le Mathématicien, No. 2.

[35] Georges SKANDALIS : *Topologie et analyse 3e année - Cours et exercices avec solutions.* Sciences Sup. Dunod, 2004.

Index

action, 12
  à droite, 12
  à gauche, 12
  classique, 62
  continue, 31
  diagonale, 14
  doublement transitive, 13
  fidèle, 13
  libre, 13
  n fois transitive, 13
  noyau de l'–, 12
  par congruence, 150
  par conjugaison, 15
  par translation, 15
  propre, 60, 141
  simplement transitive, 14
  transitive, 12
ADAMS, voir théorème d'–
adhérence, 9, 11, 27, 29, 32, 65, 95, 99, 121, 140–143, 217
adjoint (morphisme –), 175
algèbre de Lie, 283–286
algorithme de Gauss, voir pivot de Gauss
alternative de Steiner, 339
arête, 20, 118, 358, 359, 363, 366–368

BAIRE, voir espace de –, théorème de –
base incomplète, voir théorème de la base incomplète
birapport, 63, 291, 293, 296–306, 311, 315, 317, 321–323, 325, 341, 342
bloc de Jordan, voir forme normale de Jordan
BRUHAT, voir décomposition de –

caractéristique d'Euler-Poincaré, 359, 375
CARTAN, voir théorème de –
CAYLEY, voir théorème de –, transformation de –
cellule, 139, 140, 142, 255
centralisateur, 15, 16
centre
  d'un groupe, 15, 17, 78, 226, 233, 236, 237, 239, 287
cercle de $\mathbb{P}^1(\mathbb{C})$, 311, 343
CHEVALLEY, voir ordre de –
comatrice, 26, 51, 271
commutateur, 146, 284
compact, 33–35, 38, 57, 205, 206, 218, 371, 372
composante connexe, 55, 67, 69, 214, 288
composante neutre, 38, 39, 211, 214
cône isotrope, 170, 346
configuration
  de Desargues, 333
  de Miquel, 314
congruence, 150
conique, 63, 164–166, 306, 332
conjugaison
  dans un groupe, voir action par conjugaison
conjugaison harmonique, 324–326
connexe, 30, 37–39, 54, 66, 289
  par arcs, 26, 54, 71, 79, 287, 288
crochet de Lie, 283

décomposition
  de Bruhat, 16, 19, 128
  de Dunford, xvii, 98, 99, 112, 113, 122, 123, 208, 262
  de Dunford multiplicative, 208, 210
  de Frobenius, 105
  LU, 147
  polaire, xviii, 202, 211, 215, 216, 245, 288
demi-plan de Poincaré, 275

demi-plan de Poincaré, 61, 72, 74, 75, 328
dénombrable à l'infini, 35, 36, 45
dense, 26, 29, 57, 68, 84, 123, 124
DESARGUES, voir configuration de –, théorème de –
diagramme de Young, 89, 92, 96, 108, 114, 116, 118
  dual, 116
différentielle, 77, 208, 268, 271, 276, 277, 279, 281–284, 286
discriminant
  d'une forme quadratique, xviii, 62, 150, 168, 172, 186, 191, 351
  d'une matrice symétrique, xviii, 62, 150, 168, 186, 191, 351
  d'un polynôme, 29, 52, 121
distance de Hilbert, 292, 307, 320
division
  équianharmonique, 301
  harmonique, 301
drapeau, 74, 124, 128, 133, 136, 140, 141, 251, 374
droite projective, 289, 292, 293, 303, 304, 315, 318, 319, 325, 332, 337, 340, 342
dualité, 104, 338, 368, 371
DUNFORD, voir décomposition de –

ellipse de Steiner, 330
enveloppe convexe, 361, 362, 368, 369
équivalence de matrices, 2
ERLANGEN, voir programme d'Erlangen
espace
  affine, 77, 138, 255, 297
  de Baire, 36, 57
  cyclique, 102, 103
  homogène, 36, 130, 189, 202
  projectif, 40, 59, 75, 136, 141, 198, 250, 255, 256, 346
  quotient, 33, 58, 62, 80
  tangent, 268, 270, 271, 277, 281, 283, 285
EULER, voir caractéristique d'– - Poincaré
exponentielle, 70, 113, 207, 208, 210, 211, 213, 220, 221
extrémal (point –), 361, 362

face, 302, 303, 316, 317, 358–360, 363, 367–369, 372–375
facteur invariant, 104, 107, 110
fibration de Hopf, 240, 242, 312
forme
  bilinéaire, 170
  bilinéaire antisymétrique, 160, 171
  bilinéaire symétrique, 171
  de Hankel, 197
  normale, xii, 63, 73, 91, 127, 130, 139, 167, 195
  normale de Jordan, 91
  polaire, 176
  quadratique, 150, 176
formule
  de changement de base(s), 2
  de Pick, 354
  des six birapports, 301
FROBENIUS, voir décomposition de –

GAUSS, voir pivot de –, 182
grassmannienne, 40–43, 129, 130, 136, 138, 140, 142, 251, 256, 292, 293
groupe
  action de –, voir action
  affine, 49, 64, 77, 330, 361
  alterné, 236, 257, 295
  circulaire, 312
  classique, 27, 62, 71, 77, 205, 216, 249, 281
  dérivé topologique, 66
  de Lie, 330, 335
  d'isométries, 361
  d'isotropie, 153–157, 161
  de Klein, 299
  de Lie, 268, 272, 276, 281, 283
  orthochrone, 211, 215
  orthogonal, 27, 155, 159, 202, 205, 333
  parabolique, 40, 42, 74, 258

projectif, 41, 50, 250, 333, 334, 340
résoluble, 146
des similitudes, 75, 361, 373
simple, 235, 256
spécial linéaire, 27
spécial orthogonal, 27
spécial unitaire, 27
symplectique, 161
topologique, 26
unitaire, 27, 157
groupe orthogonal, 154

HAMILTON, voir quaternion, théorème de Cayley- –
HANKEL, voir forme de –
HILBERT, voir distance de –, schéma de –
homographie, 61, 72, 289, 291, 292, 294–297, 300, 305, 312, 319, 322, 325, 328, 340, 354
HOPF, voir fibration de –
hyperboloïde, 215

idéal de Lie, 285
indice de Witt, 193
invariant, xiv, 62
anallagmatique, 346
de similitude, 85, 101, 104, 107
total (ou complet), xiv
invariant de similitude, 108
inversion, 312, 313, 322, 323, 339
inversion locale, voir théorème d'inversion locale
isotrope, 158, 190, 192, 290

JORDAN, voir forme normale de –, réduite de –, théorème de –

KLEIN, voir groupe de –, programme d'Erlangen
KREIN, voir théorème de – -Milman

LEGENDRE, voir symbole de –
lemme
des noyaux, 111
des poignées de mains, 360
LIE, voir algèbre de –, crochet de –, groupe de –, idéal de –
localement compact, 35, 36, 45, 55–57, 60
LU, voir décomposition LU

matrice
échelonnée en colonnes, 130
échelonnée en lignes, 130
échelonnée réduite, 130
élémentaire, 22
antisymétrique, 160
co-échelonnée, 136
co-échelonnée réduite, 136
compagnon, 102, 103
d'une forme bilinéaire (ou quadratique), 171
diagonalisable, 28
de dilatation, 144
hermitienne, 156
hermitienne définie positive, 157
nilpotente, 62, 87
de Pauli, 228
de permutation, 144
symétrique définie positive, 154
de transvection, 144
MILMAN, voir théorème de Krein- –
MIQUEL, voir configuration de –, théorème de –

NOETHER, voir théorème de Sklolem- –
normalisateur, 15, 121, 254, 361
norme subordonnée, 52

octonion, 232, 245, 333
orbite, xi, xii, xiv, 4, 5, 8, 9, 11–13, 32, 33, 35, 60, 86, 97, 118, 131, 146, 297, 330
nilpotente, 87, 88, 92, 94, 95
semi-simple, voir classe de similitude diagonalisable
ordre
de Chevalley, 95
de dégénérescence, 10, 95, 141, 155
de dominance, 116–118

PAPPUS, voir théorème de –
partition, 84, 88, 89, 91, 93, 94, 113, 114, 126, 262, 263
  duale, 116
PAULI, voir matrice de –
PICK, voir formule de –
pinceau de droites, 304
pivot de Gauss, xvii, 18–20, 128, 130, 134, 144, 178
plan hyperbolique, xv, 190, 292
POINCARÉ, voir caractéristique d'Euler–,demi-plan de –
produit semi-direct (topologique), 43
programme d'Erlangen, xx, 291, 292
projection stéréographique, 231, 240, 309, 310, 313, 324, 355

quaternion, 224–226, 230, 240

rang, 3
réciprocité quadratique, 182, 186
réduite de Jordan, 91
régulier (sous-espace), 173
repère, 64
  affine, 64, 332
  projectif, 250, 289, 296, 336
réseau, 61, 62, 70, 75
rétraction, 47
RIEMANN, voir sphère de –

schéma de Hilbert, 80, 290
SCHLÄFLI, voir symbole de –
SCHUBERT, voir variété de –
section, 47
semi-continu, 11, 96, 218
séparé, 33, 34, 53, 55, 58–60
série génératrice, 94
SETIM, 193, 198
signature, xviii, 17, 62, 155, 157, 158, 181, 211, 215, 217–219, 274, 328, 344, 346, 351
similitude, 346, 361
  classe de –, 73, 84, 99
  classe de – diagonalisable, 84–86
  classe de – nilpotente, voir orbite nilpotente
  groupe des –s, voir groupe des similitudes
  invariant de –, voir invariant de similitude
SKOLEM, voir théorème de – -Noether
solide platonicien, 358
solide régulier, voir solide platonicien
sous-espace
  caractéristique, 111
  isotrope, 173
sous-groupe
  à un paramètre, 67, 280
  de Sylow, 254, 372
sous-groupe topologique, 66
sous-variété, 268, 270–272, 276–281, 288, 289
sphère de Riemann, 309, 310, 328, 340
stabilisateur, xviii, 6–8, 13–15, 31, 35, 60, 61, 158, 167, 251, 252, 293, 299
STEINER, voir alternative de –, ellipse de –
structure complexe, 72
submersion, voir théorème de submersion
suite exacte, 46–48, 50, 77, 214, 243, 252, 366
suite homographique, 321
SYLOW, voir sous-groupe de –
SYLVESTER, voir théorème d'inertie de –
symbole
  de Legendre, 183
  de Schläfli, 358

théorème
  d'Adams, 230
  de la base incomplète, 2, 62
  de Cartan, 268, 281
  de Cayley-Hamilton, 29
  de décomposition polaire, 202
  de Desargues, 333
  de Descartes-Euler, 359
  d'homéomorphisme, 35
  d'inertie de Sylvester, 153, 181, 182

d'inversion locale, 267, 274, 275, 283
de Jordan, 92
de Krein-Milman, 362
de Miquel, 314
du papillon, 355
de Pappus, 356
du pivot, 322
de Ptolémée, 300, 323
du rang, 3
de Skolem-Noether, 21
de submersion, 270, 278
de Sylvester, voir théorème d'inertie de Sylvester
de Witt, 159, 190
topologie quotient, 58
totalement isotrope (sous-espace), 173, 193, 197
transformation de Cayley, 327

unipotent, 113, 146, 208, 220, 221

variété de Schubert, 142

WITT, voir indice de –, théorème de –

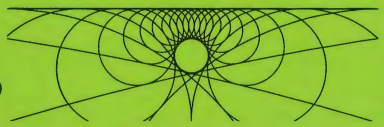
YOUNG, voir diagramme de –

Imprimé en Belgique et achevé sur les presses de SNEL Grafics, à Liège
Dépôt légal février 2013

Le présent ouvrage est avant tout une œuvre de bonheur et une invitation manifeste aux agréments mathématiques. Géométrie et groupes se donnent la main dans une valse nouvelle, aux sonorités du Programme d'Erlangen et aux modulations de la topologie et de la combinatoire.

Philippe Caldero et Jérôme Germoni nous proposent d'abord de revisiter les programmes de la licence jusqu'à l'agrégation à l'aune des actions de groupes, qui offrent un principe unificateur exceptionnel. Ces actions sont enrichies de structures variées telles la topologie ou la géométrie différentielle. À l'aide d'un nombre volontairement réduit d'outils théoriques, un plan d'étude d'une action (par la description des orbites, d'invariants, de formes normales et de l'adhérence des orbites) est mené de façon systématique dans des situations nombreuses et variées, faisant un pont entre certaines, quelque peu familières (théorème du rang), et d'autres plus sophistiquées (variétés de Schubert). La combinatoire apparaît aussi comme une version discrète de la géométrie sur les corps finis. Elle donne des applications aussi spectaculaires qu'inattendues (formule du triple produit de Jacobi comme «trace» de la théorie des matrices échelonnées, loi de réciprocité quadratique résultant de la géométrie des quadriques).

Deux auteurs, deux collègues, deux amis nous offrent là le fruit d'une collaboration heureuse, un fruit licite conçu dans le plaisir, une œuvre d'architectes, de poètes et surtout de brillants mathématiciens. Un regard nouveau et unificateur sur des thèmes classiques, particulièrement adapté à la synthèse que demande l'agrégation, avec en prime quelques perles inédites, du moins à ce niveau d'enseignement. Limitant les généralités au strict minimum, le texte offre sur deux volumes une multitude d'exemples explicites, établissant un pont entre l'algèbre linéaire, la géométrie élémentaire et des théories géométriques plus avancées.



Jérôme Germoni est maître de conférences à l'université Lyon 1 et ancien élève de la rue d'Ulm. Ancien directeur de l'IREM de Lyon, il a participé à la préparation au CAPES et à l'agrégation pendant une dizaine d'années. Ses goûts portent notamment sur les questions d'algorithmique et les problèmes de visualisation.

Philippe Caldero est maître de conférences à l'université Lyon 1 et ancien élève de l'ÉNS de Saint-Cloud. Plusieurs fois membre du jury de l'agrégation externe, il est très impliqué dans la préparation à l'agrégation interne, comme responsable et comme enseignant. Ses thèmes de recherche concernent la théorie des représentations.

Collection.– Mathématiques en devenir

ISBN 978-2-91-635231-2

Calvage & Mounet

www.calvage-et-mounet.fr

Prix : 33 €